

Journal of
Geophysical
Research

VOLUME 66 SEPTEMBER 1961 NUMBER 9

PUBLISHED BY

THE AMERICAN GEOPHYSICAL UNION

Journal of Geophysical Research

An International Scientific Publication

OFFICERS OF THE UNION

THOMAS F. MALONE, *President*
GEORGE P. WOOLLARD, *Vice President*
A. NELSON SAYRE, *General Secretary*
WALDO E. SMITH, *Executive Secretary*

OFFICERS OF THE SECTIONS

Geodesy

FLOYD W. HOUGH, *President*
CHARLES A. WHITTEN, *Vice President*
BUFORD K. MEADE, *Secretary*

Seismology

JAMES A. PEOPLES, JR., *President*
JACK E. OLIVER, *Vice President*
BENJAMIN F. HOWELL, JR., *Secretary*

Meteorology

MORRIS NEIBURGER, *President*
HENRY G. HOUGHTON, *Vice President*
WOODROW C. JACOBS, *Secretary*

Geomagnetism and Aeronomy

C. T. ELVEY, *President*
E. H. VESTINE, *Vice President*
J. HUGH NELSON, *Secretary*

Oceanography

DONALD W. PRITCHARD, *President*
ROBERT S. ARTHUR, *Vice President*
ARTHUR E. MAXWELL, *Secretary*

Volcanology, Geochemistry, and Petrology

HATTEN S. YODER, JR., *President*
EDWARD D. GOLDBERG, *Vice President*
DAVID R. WONES, *Secretary*

Hydrology

WILLIAM C. ACKERMANN, *President*
DAVID K. TODD, *Vice President*
RALPH N. WILSON, *Secretary*

Tectonophysics

LOUIS B. SLICHTER, *President*
DAVID T. GRIGGS, *Vice President*
IRIS BORG, *Secretary*

BOARD OF EDITORS

Editors: PHILIP H. ABELSON and J. A. PEOPLES

ASSOCIATE EDITORS

1959-1961

HENRI BADER	T. NAGATA
K. E. BULLEN	FRANK PRESS
CONRAD P. MOOK	A. NELSON SAYRE
WALTER H. MUNK	MERLE A. TUVE
JAMES A. VAN ALLEN	

1960-1962

JULIUS BARTELS	L. A. MANNING
V. V. BELOUSOV	TOR J. NORDENSCHELD
E. G. BOWEN	E. N. PARKER
JOHN E. CHAPPELEAR	GEORGE P. RIGSBY
G. D. GARLAND	WALTER O. ROBERTS
GORDON J. F. MACDONALD	C. N. TOWART
JAMES R. WAIT	

1961-1963

FRANKLIN I. BADGLEY	ROBERT O. REID
HENRY G. BOOKER	BRUNO ROSSI
JOSEPH W. CHAMBERLAIN	GEORGE H. SUTTON
HERBERT FRIEDMAN	DAVID K. TODD
MARK F. MEIER	VICTOR VACQUIER

ARTHUR H. WAYNICK

The Editors of the *Journal of Geophysical Research* welcome original scientific contributions on the physics of the earth and its environment.

Manuscripts should be submitted in triplicate to J. A. Peoples, Jr., Department of Geology, University of Kansas, Lawrence, Kansas. Authors' institutions in the United States or Canada, are requested to pay a publication charge of \$25 per page, which, if honored, entitles them to 100 free reprints.

Subscriptions to the *Journal of Geophysical Research* and *Transactions, AGU*, are included in membership dues.

Nonmember subscriptions, *Journal of Geophysical Research*, \$30 for back volume of 1959, \$45 for back volume of 1960, \$6 for this issue; \$25 for the calendar year 1961.

Nonmember subscriptions, *Transactions, AGU*, \$4 per calendar year, \$1.25 per copy.

Subscriptions, renewals, and orders for back numbers should be addressed to American Geophysical Union, 1515 Massachusetts Ave., Northwest, Washington, D. C. Suggestions to authors are available on request.

Advertising Representative: Howland and Howland, Inc., 230 Park Ave., New York 17, N. Y.

Since January 1959 (Vol. 64, No. 1) the *Journal of Geophysical Research* has been published monthly by the American Geophysical Union, the U. S. National Committee of the International Union of Geodesy and Geophysics, organized under the National Academy of Sciences-National Research Council as the U. S. national adhering body. Publication of this journal is supported by the National Science Foundation and the Carnegie Institution of Washington. The new monthly combines the type of scientific material formerly published in the bi-monthly *Transactions, American Geophysical Union*, and the quarterly *Journal of Geophysical Research*. The *Transactions, American Geophysical Union*, will continue as a quarterly publication for Union business and items of interest to members of the Union.

Published monthly by the American Geophysical Union from 1407 Sherwood Avenue, Richmond, Virginia. Second class postage paid at Richmond, Virginia.

... The gravimetric method is certainly the most powerful branch of geodesy because *it can relate the different continents to the same world geodetic system*; it is the only method capable of giving the detailed shape of the geoid not only on the continents but also over the oceanic areas. It . . . *

* W. A. Heiskanen in "Latest Achievement of Physical Geodesy" (J. of Geophys. Research 65, p. 2827-2936, Sept. 1960)



GRAVIMETER SURVEYS OVER LARGE DISTANCES

the ASKANIA GRAVIMETER Gs-12

entire gravitation force acting on measuring system (M.g.) is nearly constant. Any variations in gravity are compensated through a change in mass of the measuring system by adding or removing small balls of an accurately established mass so that the system is largely unaffected by the measuring characteristics

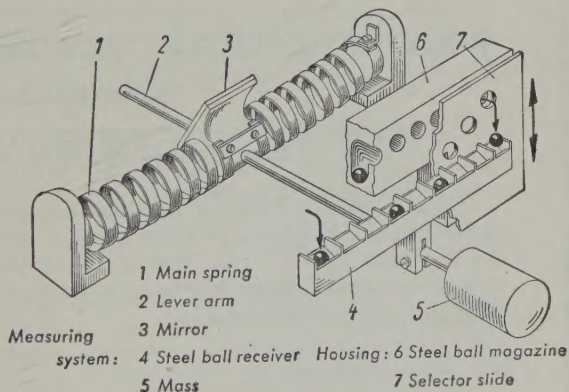
measuring range about 2,000

measuring accuracy 0.02 mgal

gravimetric surveys of the oceans rely on the ASKANIA Seagravimeter Gss-2 after Graf.

AVAILABLE: Earth Tide Recording Equipment, Magnetic Instruments, etc.

for detailed information



ASKANIA-WERKE

Division of Continental Elektroindustrie AG

Branch

4913 Cordell Avenue

Bethesda 14, Maryland

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers

SPRENGNETHER'S DIRECT WRITING VISUAL RECORDER PROVIDES 24 HOUR REGISTRATION.

For a moderate initial investment, seismological laboratories can obtain this superior drum-type recorder. Its advantages include continuous registration and easy visual access to all information on the recorder. It also requires less storage space for records than a tape recorder.

Drum is completely enclosed to protect against dust or accidental damage. For greatest possible convenience in changing records, the large, curved plastic cover can be fully opened.

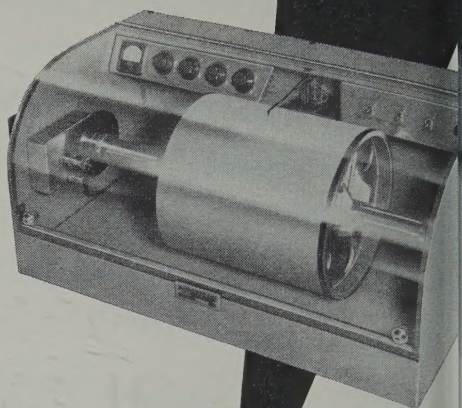
The pen drive galvanometer is a high torque, frictionless, torsion-type moving coil system. It operates in the field of large Alnico V magnet with special pole pieces and core.

Pen-drive is rigidly mounted on inside of back panel of recorder. Ink-well is mounted on the axis of moving coil to eliminate inking problems.

Recorder box is of heavy steel and has a beautiful, baked enamel finish.

SPECIFICATIONS:

Dimensions: 31" long, 16" wide, 15" high (not including motor housing) — Weight: 90 lbs. net — Drum Speed: 30 or 60 mm/min. — Translation Rate: 2.5 or 5 mm/rev. Power Requirements: 110 V. or 240 V. AC, 50 or 60 cy. Paper Size: 36" x 12" — Pen Galvanometer Sensitivity: .1 milliamperes/mm.



Series VR-40-0

Shown with controls as used with our VR-30-A Amplifier.

Write for complete technical information today.

Internationally Known Mfrs. of Seismological, Geophysical Instr
W. F. SPRENGNETHER INSTRUMENT COMPANY
4567 SWAN AVE. • ST. LOUIS 10

BULLETIN (IZVESTIYA), ACADEMY OF SCIENCES, U.S.S.R. GEOPHYSICS SERIES

Subscriptions for 1961 series now available

This monthly Russian publication, perhaps the leading journal of Geophysics of the U.S.S.R., is being translated and published in an English edition for the year 1961 by the American Geophysical Union. The twelve numbers in Russian cover about 2000 pages. Published with the aid of a grant from the National Science Foundation.

Send subscriptions now to

AMERICAN GEOPHYSICAL UNION
1515 Massachusetts Avenue, N.W.
Washington 5, D. C., U.S.A.

Subscription rates: \$25.00 for the volume of 12 numbers (\$20.00 to individual members of AGU subscribing for personal use)
Numbers will be mailed as issued.

The English edition of this publication for 1957 has been translated and published for the American Geophysical Union by Pergamon Press. This volume may be ordered through the American Geophysical Union at a price of \$25.00. The 1958, 1959, and 1960 series are available at a price of \$25.00 for each volume of 12 numbers. Titles and authors of the papers contained in the series have been published in recent issues of the *Transactions, AGU*.

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertisers



IONOSPHERIC PROPAGATION AND HF COMMUNICATIONS

If you have the background, the imagination and the desire to contribute to important programs in these fields, you are invited to join a carefully selected team of outstanding scientists and engineers now contributing significantly to current knowledge through advanced research.

Our present needs are for:

SENIOR IONOSPHERIC PHYSICISTS

Ph.D. preferred, with several years' experience in the study of Ionospheric phenomena. Should be familiar with present knowledge of upper atmosphere physics and possess an understanding of current programs using rockets and satellites for studies in F-region and beyond. Qualified individuals with supervisory abilities will have an exceptional opportunity to assume project leadership duties on HF projects already under way involving F-layer propagation studies backed by a substantial experimental program.

SENIOR DEVELOPMENT PHYSICISTS

Advanced degree in Physics or E.E. preferred. Must be familiar with latest techniques in the design of advanced HF receivers and transmitters and possess working knowledge of modern HF networks employing ferrites and metallic tape cores. Strong theoretical background in modern linear circuit theory desired. Will carry out laboratory development and implementation of new HF communications systems.

SENIOR ELECTRONIC ENGINEERS

Advanced degree in E.E. preferred. Must be familiar with conventional pulse circuit designs and applications. Technical background should include substantial experience in data process and data recovery systems using both analog and digital techniques. Knowledge of principles and application of modern information theory including correlation techniques helpful. Will be responsible for the design of sub-systems.

These programs are being conducted at our ELECTRO-PHYSICS LABORATORIES in the suburban Washington, D.C. area, ideally located from the viewpoint of advanced study which may be conducted at one of several nearby universities; for readily available housing in pleasant residential neighborhoods; and for the general amenities of living offered by this important Metropolitan center.

All qualified applicants will receive consideration for employment without regard to race, creed, color or national origin.

*For a prompt reply to your inquiry,
please forward resume in confidence to:*

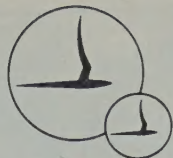
W. T. WHELAN
Director of Research & Development

ACF ELECTRONICS
DIVISION

ACF INDUSTRIES

HYATTSVILLE, MARYLAND

HEAD ATMOSPHERIC PHYSICS SECTION



As a part of important programs in the atmospheric sciences now under way at CAL, our scientists are seeking answers to stimulating questions in the fields of:

CLOUD PHYSICS

RADAR METEOROLOGY

WEATHER MODIFICATION

DYNAMIC METEOROLOGY

ATMOSPHERIC ELECTRICITY

We require a man qualified to assume overall leadership of these programs. It is an excellent opportunity for an Atmospheric Physicist to take over the senior position in a well established and growing research program. University associated atmosphere. Ample opportunity to develop research programs in areas of prime interest to you. Excellent potential for growth.

For additional information, address

C. C. Rentschler
CORNELL AERONAUTICAL
LABORATORY, INC.
of Cornell University
Buffalo 21, New York

All qualified applicants will receive consideration for employment without regard to race, creed, color or national origin.

ARE YOU A MEMBER of AGU?

If you are a reader or user of either of the publications noted below, you should look into this matter.

Members regularly receive the monthly

**Journal of
Geophysical Research**

the quarterly

**Transactions
American Geophysical Union**

Members are entitled to special discounts on other publications of AGU such as

- **Geophysical Monograph Series** (Nos. 1 to 5 available)
- **Izvestiya of the Academy of Sciences, USSR; Geophysics Series** for the years 1957, 1958, 1959 and 1960
- **Geodesy and Cartography, USA** 1959

- Annual meeting in Washington late April or early May each year
- Regional meeting in the Pacific Northwest in autumn
- Regional meeting in the Pacific Southwest in winter
- National Western Meeting, December 1961

For application forms and other information regarding membership, meetings, and other matters, write:

AMERICAN GEOPHYSICAL UNION
1515 Massachusetts Ave., N.W.
Washington 5, D. C.

(An application form is also inserted in the closing pages of this issue.)

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertiser.

depth



The new Sperry Rand Research Center has been established to pursue a specific objective—the advancement of man's knowledge and understanding through basic research in the broad and diversified scientific fields that are of potential long-range interest to the many technical activities of the Sperry Rand Corporation.

To achieve this ambitious goal, many significant investigations must be pursued to great depths.

Investigations in depth are vitally important to Sperry Rand in the following areas:

1. **Solid State Physics** including theoretical and experimental studies of magnetic, electrical, optical, chemical, and mechanical properties.
2. **Information Processing** including such areas as communication, information and coding theory; automatic control; modulation, generation and detection techniques; human factors.
3. **Plasma Physics** including electromagnetic wave interactions, reentry problems, propagation studies.
4. **Applied Mathematics** including probability and statistics, discrete mathematics, solid and fluid mechanics.
5. **Earth Sciences** including geophysics, physical oceanography, atmospheric physics, and special topics in Life Sciences.

At the present time professional openings are available to senior scientists with experience in:

- Semiconductor Physics
- Magnetic Resonance Phenomena
- Solid State Theory
- Physical Chemistry
- Communication Theory
- Signal Processing
- Theory & Application
- Automatic Control Theory
- Plasma Physics
- Wave Propagation
- Reentry Physics
- Stochastic Processes
- Game Theory
- Fluid & Solid Mechanics
- Operations Research
- Geophysics
- Oceanography
- Atmospheric Physics

Inquiries may be directed in complete confidence to:
Mr. Frederick M. Swope, Jr.,
Sperry Rand Research Center,
North Road,
Sudbury, Massachusetts

an important
word
at the new

SPERRY RAND RESEARCH CENTER
SUDBURY, MASSACHUSETTS

GEOPHYSICAL MONOGRAPH SERIES

AMERICAN GEOPHYSICAL UNION

1515 MASSACHUSETTS AVENUE, N.W.

WASHINGTON 5, D. C., U.S.A.

Antarctica in the International Geophysical Year—Geophysical Monograph No. 1 (Publication No. 462, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 56-60071; 133 pp. and large folded map of the Antarctic, 1956, 7" x 10", \$6.00. Contains 16 pages by various American authorities on the Antarctic under the headings: General, Geographic and Meteorological, Geological and Structural, Upper Atmospheric Physics, and Flora and Fauna. Map (41" x 41") compiled by the American Geographical Society. Introduction by L. M. Gould.

Geophysics and the IGY—Geophysical Monograph No. 2 (Publication No. 590, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 58-60035; 210 pp., 1958, 7" x 10", \$8.00. Contains 30 papers by leading American authorities under the headings: Upper Atmospheric Physics, The Lower Atmosphere and the Earth, and The Polar Regions. Preface by Joseph Kaplan.

Atmospheric Chemistry of Chlorine and Sulfur Compounds—Geophysical Monograph No. 3 (Publication No. 652, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 59-60039; 129 pp., 1959, 7" x 10", \$5.50. Based on a symposium held jointly with the Robert A. Taft Sanitary Engineering Center of the U. S. Public Health Service in Cincinnati in November, 1957. Contains 23 papers (some as summaries) with discussion. Preface by James P. Lodge, Jr.

Contemporary Geodesy—Geophysical Monograph No. 4 (Publication No. 708, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 59-60065; 96 pp., 7" x 10", 1959, \$5.50. Based on a Conference held at Cambridge, Massachusetts, in December 1958 jointly by the AGU with the Smithsonian Astrophysical Observatory and the Harvard College Observatory. Contains 14 papers by leading authorities, with verbatim discussions on topics ranging from classical geodesy to trilateration by underwater sound to space navigation in the solar system. Edited by Charles A. Whitten and Kenneth H. Drummond.

Physics of Precipitation—Geophysical Monograph No. 5 (Publication No. 746, National Academy of Sciences—National Research Council); Library of Congress Catalogue Card No. 60-60010; 435 pp., 7" x 10", 1960, \$12.50. Based on a Conference held at Woods Hole, Massachusetts, in June 1959. Contains 48 papers by leading authorities, with verbatim discussions on topics ranging from planetary-scale phenomena to microanalysis including hail formation and precipitation control. Edited by Helmut Weickmann.

Postage is to be added to prices shown unless payment accompanies order. Quantity discounts (count each Monograph separately): 5-19 copies, 10%; 20-49 copies, 15%; 50 or more copies, 20%.

Purchase Order

TO AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Avenue, N.W., Washington 5, D. C., U.S.A.

Please enter our order for the following:

_____ copies of Geophysical Monograph No. 1, at \$6.00 *	\$ _____
_____ copies of Geophysical Monograph No. 2, at \$8.00 *	\$ _____
_____ copies of Geophysical Monograph No. 3, at \$5.50 *	\$ _____
_____ copies of Geophysical Monograph No. 4, at \$5.50 *	\$ _____
_____ copies of Geophysical Monograph No. 5, at \$12.50 *	\$ _____

- ☐ Payment of \$ _____ is enclosed.
- ☐ Please send invoice, adding postage charges.
- ☐ Enter our standing order for _____ copies of subsequent Geophysical Monographs at the special prepublication rates, e.g., prepublication rate for Monograph No. 4 for non-members was \$4.00, payment in advance, or \$4.75 (plus postage) on invoice.

* List price is net for quantities up to four; see above for discounts on quantity purchases. Special discounts to members.

Typed name _____ Signature _____

Address _____

Please mention JOURNAL OF GEOPHYSICAL RESEARCH, when writing to advertiser.

Journal of GEOPHYSICAL RESEARCH

VOLUME 66

SEPTEMBER 1961

NO. 9

Dynamics and Structure of the Outer Radiation Belt

C. Y. FAN,¹ P. MEYER, AND J. A. SIMPSON

*Enrico Fermi Institute for Nuclear Studies
University of Chicago, Chicago 37, Illinois*

Abstract. From an analysis of electron measurements in the Explorer VI satellite (August 7–October 6, 1959) four time-dependent parameters which characterize the outer electron belt have been investigated. They are: (1) the equatorial electron intensity I_0 ; (2) the equatorial range from the center R_0 of the peak intensity; (3) the electron-density distribution along a line of force through the intensity peak; and (4) a measure of the change in electron spectrum with time. These parameters, along with measurements of magnetic field intensity, make it possible to study the origin of the changes in electron intensity and distribution which are known to occur in the outer belt. Several magnetic storms occurred during the observation on Explorer VI. Within the sequence of changes in the outer belt induced by these geomagnetic storms, there are some changes of the parameters which are accounted for only by invoking an irreversible energy gain or loss within the outer belt. The energy gain process appears to be through irreversible local acceleration of electrons. The energy loss process leads to a stable mirror-point distribution characteristic of the periods between geomagnetic storms. The time intervals within which each of these processes is operative are identified. Reversible processes are possibly the cause for other changes. The foregoing analysis rests upon the proof given in this paper that the outer-belt peak intensity coincides over a wide range of geomagnetic latitudes with magnetic field lines of force in the centered dipole approximation. Consequently, the measured electron-intensity maximum is used as a 'tracer' of the geomagnetic field lines of force for analyzing changes in the outer belt with time. It is shown that even during geomagnetic storms the trace of the electron-intensity maximum followed a centered dipole line of force. This indicates that at all times the particle-energy density in the radiation belt is much less than the energy density of the magnetic field in the region. The electron fluxes, high-energy proton fluxes, and possible electron spectra are investigated. Two distinct peaks of electron intensity are identified to persist in the outer belt for about 2 months, and it is shown that these peak distributions undergo radial motion during geomagnetic disturbances.

1. INTRODUCTION

The first measurements of charged particles within the geomagnetic field were made at Iowa [Van Allen, McIlwain, and Ludwig, 1959] and by the U. S. S. R. groups [Vernov, Gerasimov, Vakulov, and Logachev, 1959a]. Subsequent observations proved the existence of an inner belt composed of energetic protons and alpha particles, and an outer radiation belt dominated by energetic electrons. Single traversals of the outer belt made at different times led to the

conclusion that the characteristics of this outer belt changed with time and solar activity. The origin of the inner belt protons is satisfactorily accounted for by the decay products of fast neutrons escaping from the terrestrial atmosphere. Although the origin of the outer belt at first was ascribed to electrons accelerated at the sun and subsequently trapped in the geomagnetic field, this hypothesis foundered on both the basis of physical principles and through experimental evidence, including results reported in this paper. Therefore, if electrons are to be introduced from the sun, or other sources at very low energy, it is required that they be accelerated in the geomagnetic field to yield the

¹ Present address: Laboratories for Applied Sciences, University of Chicago, Chicago, Illinois.

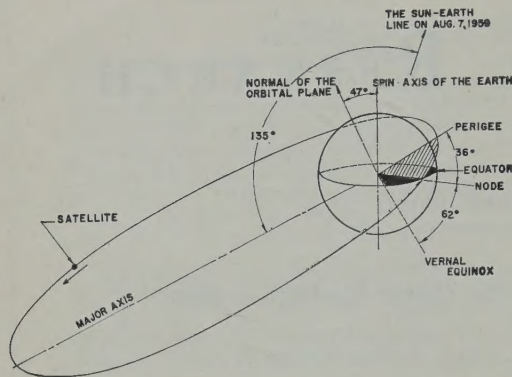


Fig. 1. Perspective drawing of the Explorer VI satellite orbit. The orbital data are as follows: period, 12 hours 48 minutes; apogee, 48,788 km; perigee, 6,626 km; Geographic inclination, 47° ; Tilt of orbit, 38° with respect to ecliptic plane.

persistent, variable flux of energetic electrons observed in the outer radiation belt. On the other hand, if the outer-belt electrons are ambient electrons, or electrons from neutron decay [Hess, 1960; Dessler and Karplus, 1960], then the changes in electron intensity and spatial distribution must also arise from magnetic field changes.

Therefore we conclude that a study of the changes in the electron outer belt brought about by changes in the geomagnetic field may lead us to the origin of these trapped electrons.

The energy that produces the variations of the external geomagnetic field comes from solar events, such as the solar flares, which emit ionized plasma leading to the classical geomagnetic storm.

There are a variety of ways in which some of these changes in energy and distribution of trapped electrons may be brought about [e.g., Northrop and Teller, 1960; Parker, 1961a]. Either rapid field changes including hydromagnetic wave propagation lead to the violation of one or more of the adiabatic invariants, and hence to irreversible acceleration or loss of energy in the belt, or slow field changes can dominate to produce reversible acceleration and consequent changes in the particle distributions in space.

The main purpose of the present investigation is to decide whether reversible processes determine the changing character of the outer electron belt or whether irreversible processes must be invoked. The experiment, therefore, requires frequent, sequential traversals of the outer

electron region at intermediate and low latitudes for extended periods of time under quiet conditions as well as during geomagnetic storms. An investigation of this kind was first undertaken by the Explorer VI satellite launched August 7, 1959, in the highly elliptic orbit shown in Figure 1. This satellite carried a group of charged-particle radiation detectors having a wide dynamic range and variety of energy responses, as well as a magnetometer. The types of radiation detection instruments are given in Appendix A, Table 1, along with a list of space vehicles which at this writing have passed through the outer belt at low latitudes (Appendix A, Table 2). The measurements were obtained under fortunate circumstances, since the magnetic field remained quiescent for a week before the commencement of two geomagnetic storms closely spaced in time. All the data within the period August 7 through October 1959, following which time the transmission ceased to operate.

A preliminary account of our experiment [Fan, Meyer, and Simpson, 1960] has shown several new properties of the trapped electrons in the outer belt, such as the persistent existence of a double belt structure in the outer zone, radial motions of these peak intensity distributions with time. We also observed the large-time-dependent changes of intensity and particle distributions following geomagnetic storms which have also been reported by other investigators with apparatus on Explorer VI [Arnoldy, Leman, and Winckler, 1960a; Rosen, Farley, Sonett, 1960]. However, our detailed study of conclusions regarding the outer-belt electrons have awaited the full reduction of data, the availability of correlating solar geophysical data, and improved orbit calculations for Explorer VI.

To prove whether there exist irreversible changes or only reversible changes, our investigation centers on the identification and measurement of parameters that describe the characteristics of the outer belt and their changes with time. We first show that a centered dipole magnetic field is a good approximation for the description of trapped electron guiding centers in the outer belt, even during periods of geomagnetic storms. Thus, the positions of intensity maxima in the outer belt becomes 'tracers' which we follow changes in position, intensity, or other parameters as a function of time.

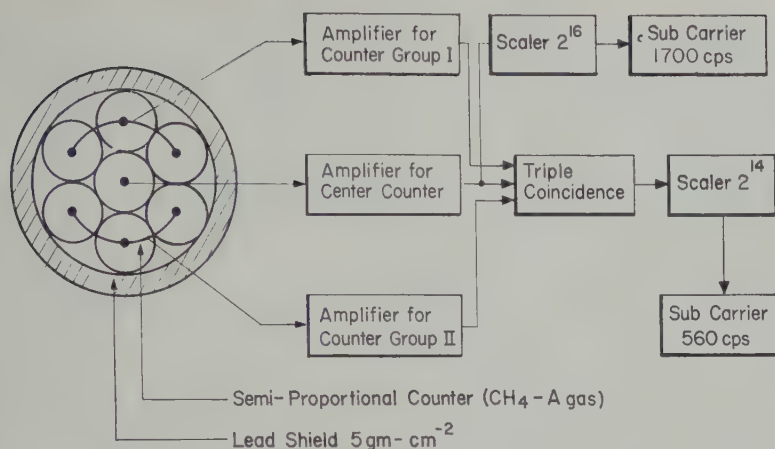


Fig. 2. Block diagram of instrumentation in Explorer VI used for the analysis of the outer electron belt.

possible to follow systematically through the progress of a geomagnetic storm the changes in parameters such as the equatorial range R_0 of the outer belt, or the equatorial intensity maximum I_0 . To investigate reversible and irreversible processes we also need to know the electron mirror-point distribution along a magnetic line of force and the changes in this distribution with time. Having proved that the centered ellipse approximation describes the lines of force in the outer belt, we are able to construct experimental curves of electron intensity as a function of position along magnetic lines of force to relate the changes in their distributions to geomagnetic storms.

The analysis of bremsstrahlung intensity rests on some knowledge of the changes in electron spectrum with time and position in the geomagnetic field. Through the introduction of a spectrometer that reveals when the electron spectrum changes as a function of position or time, we can detect spectral changes during magnetic storms, changes in mirror-point distributions, etc. Measurements show that there is a concentration of electrons at the equator during some geomagnetic storms.

Our results lead to the identification of two reversible processes among all the changes in the electron outer belt during geomagnetic storms. One is an energy or particle gain; the other is an energy or particle loss. Although not proved, it also appears unlikely that the outer belt electrons are solar electrons which

arrive with subdetection energies to undergo local postacceleration. It is probable that the irreversible processes of gain and loss operate on ambient electrons and beta decay electrons. Some conditions are established for developing specific models of the electron outer belt.

2. CHARACTERISTICS OF THE DETECTOR SYSTEM

In order to separate high-energy protons (> 75 Mev) and electrons ($E > 13$ Mev) from bremsstrahlung produced by lower energy electrons, a triple coincidence telescope consisting of seven methane-argon filled gas counters was used for the detectors (Fig. 2). These counters operated in the semiproportional range with a dead-time of 0.8 microsecond. This makes it possible to distinguish the cosmic ray flux from the trapped radiation in the outer radiation belt and to separate the bremsstrahlung production from energetic protons in the inner radiation belt. Groups of these counters were combined as shown in Figure 2 to form a wide-angle, triple-coincidence counter telescope, so that the accidental rate would be negligible even for the highest fluxes expected in the outer belt. The wide-angle was chosen to avoid the variation of counting rates due to any possible anisotropy of the radiation in the geomagnetic field. The lead shield surrounds the triple-coincidence telescope for two reasons: (1) It provides a threshold for the minimum energy protons detectable by the triple-coincidence telescope (75 Mev for protons and ≥ 13 Mev for electrons). (2) The 5 g-cm^{-2}

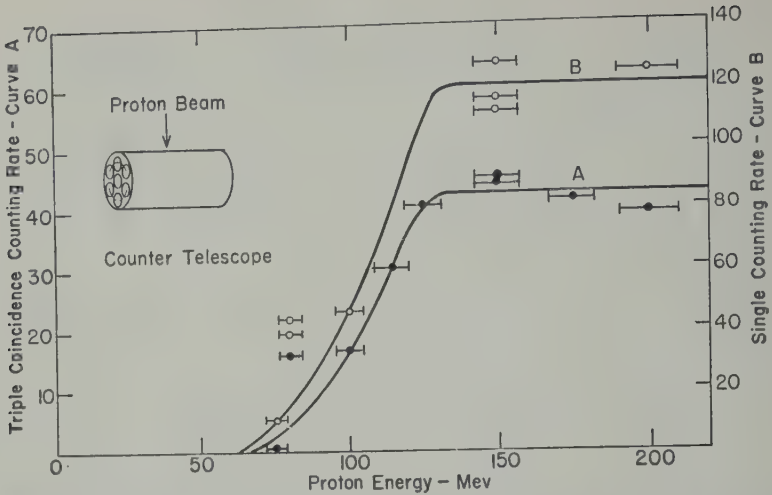


Fig. 3. Detector response for protons from the synchrocyclotron beam of the University of Chicago.

of lead reduces the total counting rate in each of the individual counters to such a level that even for the highest fluxes attainable in the outer radiation belt, only minor corrections due to the dead-time were needed, and never more than 2 per cent.

The information from the outputs of the two scalers in Figure 2 were transmitted to ground receiving stations through two subcarrier channels in the Explorer VI satellite.

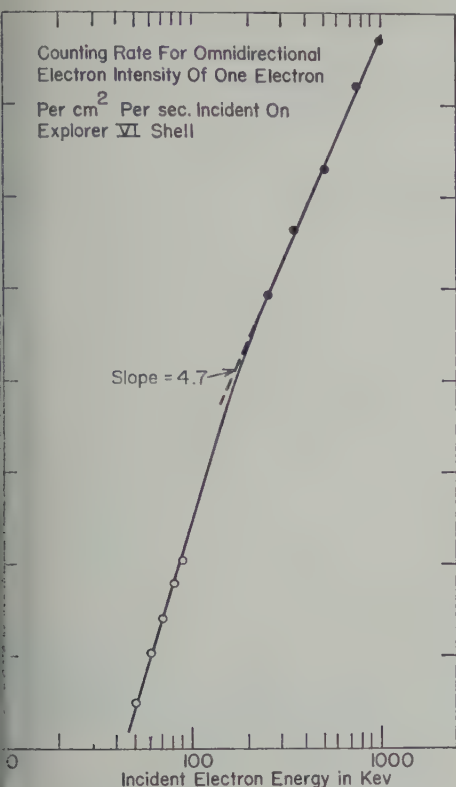
The response of this detector system to protons was studied with the University of Chicago synchrocyclotron proton beam. The results are shown in Figure 3 for the detector axis perpendicular to the proton beam over an energy range of 50 to 250 Mev. In measuring relatively flat spectra like the cosmic radiation and the inner-belt protons [Freden and White, 1960] this detector possesses a sharp cutoff at the mean energy of 100 Mev. For solar flare protons where the spectra may be as steep as $1/E^2$, the gradual change in sensitivity in the region from 75 to 100 Mev, as shown in Figure 3, must be taken into account.

The noncoincident high counting rates of the single counter during the passage through the outer belt must be due exclusively to geomagnetically trapped electrons which produce bremsstrahlung either in the satellite shell, interior objects, or in the surrounding lead shield. With an unknown energy spectrum it is obviously impossible to relate the count rate of

the single counter to the flux of particles or the energy flux carried by the electrons incident on the vehicle. However, with proper calibration we shall show in section 11 that significant limits on the particle flux and on the particle energy spectra may be derived.

The response of this bremsstrahlung detector to monoenergetic electrons was studied in the laboratory using the electron beam from a Van de Graaff generators (High-Voltage Engineering Company and the General Electric Company) and a low energy X-ray source at the University of Minnesota. The completely assembled payload of Explorer VI, with a mass distribution identical to the flight unit, was mounted so that it could be rotated about two mutually perpendicular axes for exposing any portion of its surface to direct electron bombardment. (The electron flux in the energy range from 0.1 keV to 1 Mev in discrete energy intervals irradiated the payload from various directions so that an averaged response of the counter to isotropic radiation could be deduced. This experiment differs from an exposure to particles in space in two ways: there was air present inside the payload; the electron beam for the exposure was unidirectional instead of isotropic. Additional measurements were obtained inside the vacuum system using a bremsstrahlung detector with an aluminum shell simulating the satellite shell.)

Since the shell of Explorer VI payload is



4. Calibration curve for the center counter in Figure 2. These electron data were obtained with the detector in the duplicate Explorer shell.

aluminum 170 mg/cm² thick, all electrons of energies below 500 kev are stopped in the shell. Therefore the response of our radiation detector to electrons below 500 kev is due entirely to bremsstrahlung produced in the payload shell. This was verified by showing that the counting rate for an equivalent electron flux was proportional to the inverse square of the distance between the detector and the irradiated area on the shell. The angular distribution of X rays from the satellite appears to be sufficiently isotropic to use the inverse square law for those portions of the payload where no appreciable shielding material is located between the payload shell and our detector. The assumption of isotropy for the X rays produced in the shell is an even better approximation in the case of omnidirectional incidence of electrons in space.

in the equatorial plane of the vehicle where

the instruments and associated equipment are located, the experimental count rates for the electron flux incident at various payload longitudes were used to compute the bremsstrahlung contribution. The openings between equipment and electronic parts in the payload were studied, and the calibration takes into account factors that could influence the count rate from low energy X rays.

From these measurements we wish to obtain the flux of monoenergetic electrons isotropically incident upon the satellite from the counting rate of our bremsstrahlung detector. To this end we integrated over the response curve in the equatorial plane, and added the upper and lower hemisphere integrations of the payload shell separately. Thus we obtained the response curve in counting rate per electron per square centimeter bombarding the shell as shown in Figure 4. The portion of the curve between 200 and 1000 kev can be expressed as a power law $E^{4.7}$.

The performance of our detector system for separating trapped protons and cosmic radiation from bremsstrahlung is clearly shown in the Explorer VI measurements of Figure 5 which represents a traversal through the inner and outer radiation belts at high geomagnetic latitudes.² The triple-coincidence counting rate rises to a maximum value at 8500 km representing the peak of the inner radiation belt for protons > 100 Mev, and decreases to a cosmic-ray background at the radial distance of 9500 km and beyond.

For counting penetrating particles in a single counter the calculated omnidirectional cross section is 5.5 cm², which is 1.95 times the cross section for triple coincidences measured with Explorer VI, and 1.93 with Pioneer V. These measurements were obtained by using the omnidirectional cosmic-ray intensity at great distances from the earth with no bremsstrahlung present. Any increase in this ratio signifies bremsstrahlung detected by the single counter. For example, in the outer belt the bremsstrahlung rate reaches 1.8×10^4 counts/sec compared with a triples count rate of 5 to 6 counts/sec from cosmic radiation.

² During each period of satellite motion there are two passes through the Van Allen belts; we define an ingoing or outgoing traversal of the trapped radiation as a pass, numbered consecutively from launch. The dates of the consecutively numbered passes are given in Appendix A, Table 3.

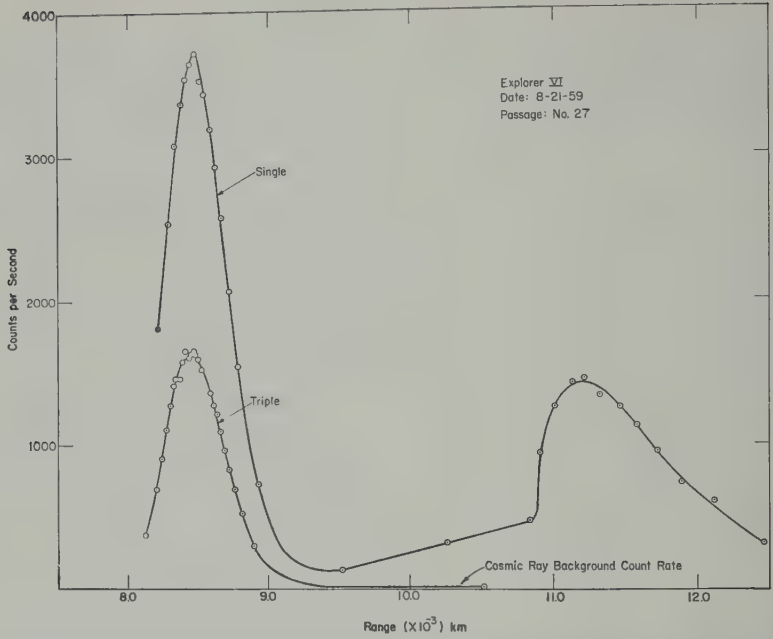


Fig. 5. Counting rate data along the satellite trajectory (pass 27) showing the existence of the high-energy proton region P_1 (inner belt) and the E_3 peak from electrons in the outer belt. Note that the triple-coincidence (proton) count rate beyond P_1 is due to cosmic radiation alone.

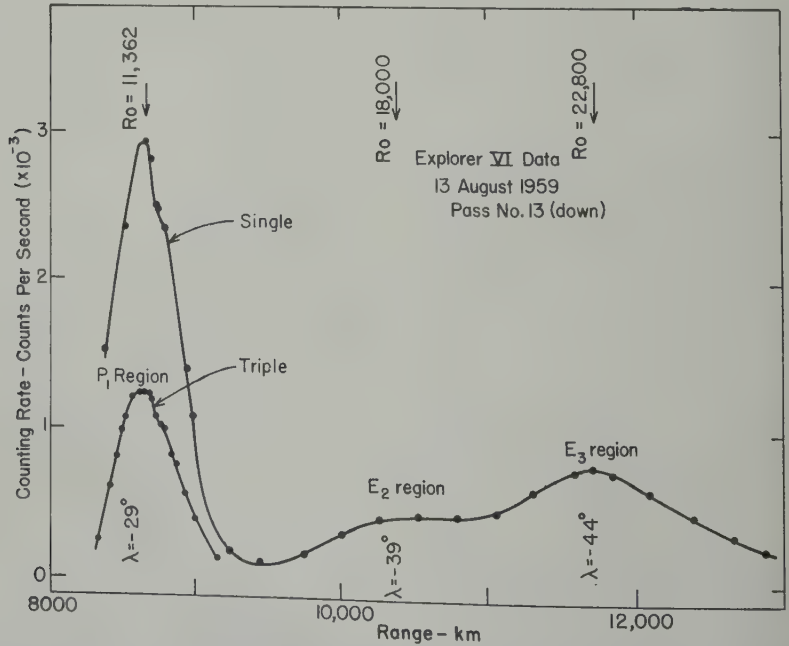
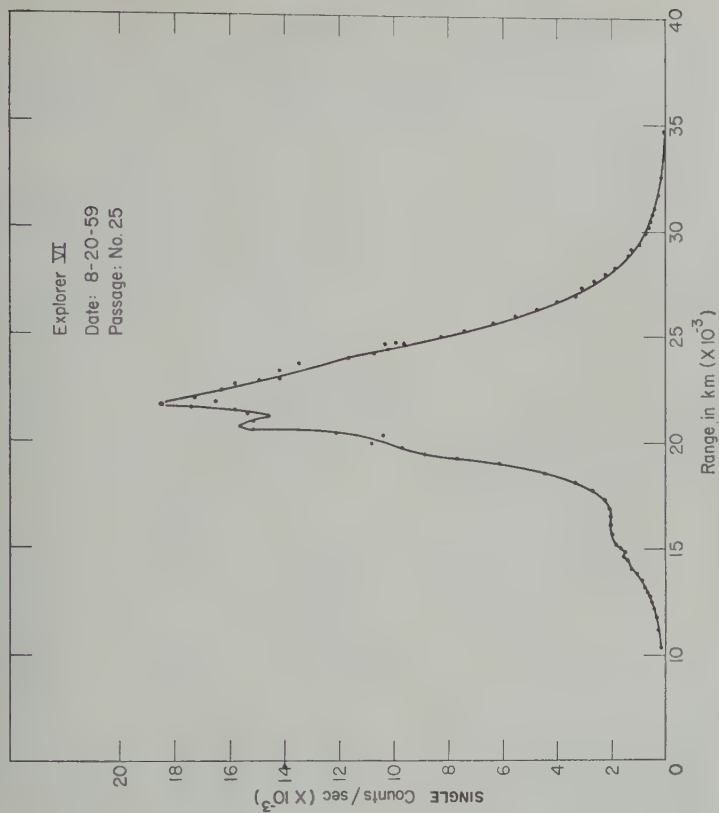
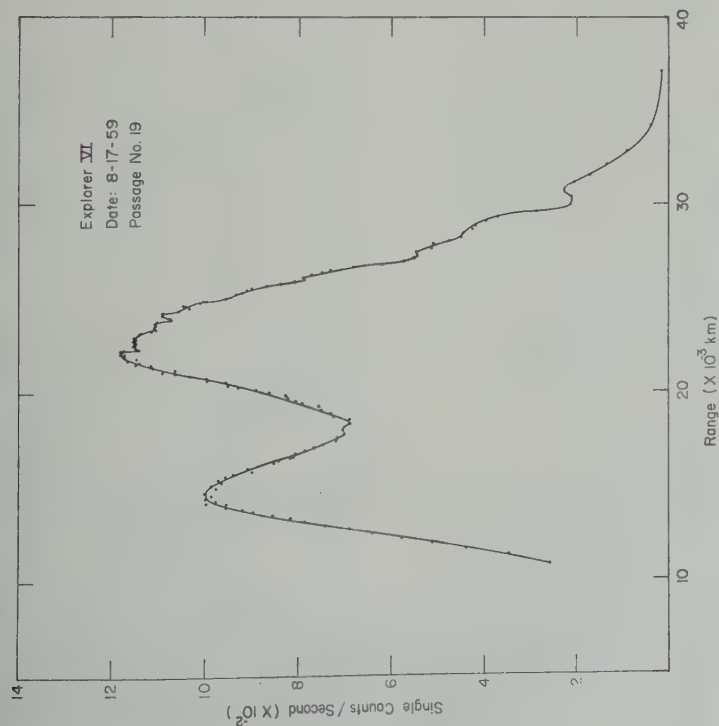


Fig. 6. Counting rate data along the satellite trajectory (pass 13) showing the existence of regions P_1 (inner belt), E_2 , and E_3 (outer belt). Statistical errors are approximately the size of individual data points.



Figs. 7a, b. Counting rate from bremsstrahlung as a function of range along the satellite trajectory. Electron peaks E_2 and E_3 are shown. All data points determining the curves for each pass are shown here. Statistical errors are less than ± 1 per cent.

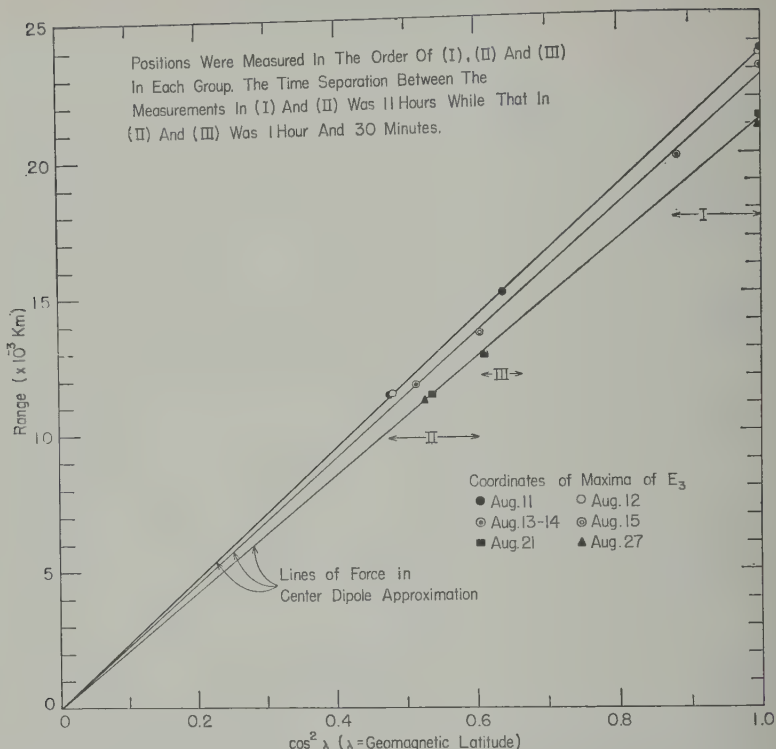


Fig. 8. Proof that the outer belt maximum E_3 (max) lies on magnetic lines of force in the center dipole approximation even for periods of geomagnetic storms.

tion. In Figure 5 it is seen that detectable bremsstrahlung sets in on the outer slope of the inner radiation belt and increases until reaching a peak intensity of the outer belt. These measurements demonstrate the separation of the detection of bremsstrahlung from energetic nucleonic particles. It is also interesting to note that the triple-coincidence detector has measured untrapped solar flare protons while it was in the outer radiation belt where the highest singles count rate was 6500 counts/sec [Fan, Meyer, and Simpson, 1960].

3. THE EXISTENCE OF TWO ELECTRON INTENSITY PEAKS

Typical data recorded during passes through the trapped radiation regions are shown in Figure 6 and Figures 7a, b. They immediately reveal the existence of the two electron intensity peaks reported in 1959 [Fan, Meyer, and Simpson, 1960]. Using an ion chamber and Geiger counter, the Minnesota group have confirmed the existence of this double-electron-peak structure

[Arnoldy, Hoffman, and Winckler, 1960]. Our data show throughout the 8 weeks Explorer VI recordings that a double peak is a persistent characteristic, and since a reexamination of earlier space-shot data suggests the possible existence of this intermediate peak even over greater periods of time, we suggest that two separated electron intensity maxima may be a persistent feature of the radiation belt. Therefore this paper uses the nomenclature, proposed earlier, designating the subscripts 1, 2, 3 the peak intensity at successive radial distance, and indicating by the letter P or E whether we observe protons or electrons: i.e., the electron regions E_1 , E_2 , and E_3 are the outer Van Allen belt; thus the peaks in Figure 6 are identified as P_1 , E_2 , and E_3 .

The two peaks in Figures 7a and 7b are the E_2 and E_3 regions. To reduce the complexity of VI observations to a quantitative study, it is necessary to trace the identity of these peaks through intensity changes or positional changes of the outer belt with time before, during, and after

geomagnetic disturbances. Therefore, we defined the maximum intensity of E_2 and maximum intensity of E_3 as 'signatures' of E_2 and E_3 regions, and we use these maxima as the E_2 and E_3 electron distributions in the geomagnetic field both in space and time. For example, for a given peak distribution at the time during times of quiescent geomagnetic disturbances, the peak intensity is to be assigned to a specific line of force of the dipole field at R_0 , as we show in the next section. With the assumptions that the guiding center approximation holds for the trapped electrons, and that the energy spectrum and the pitch-angle distribution do not drastically change in the neighborhood of the line of force through E_3 , we find that the maximum intensity of electrons will be along a line of force over a wide range of latitudes for which the electron is trapped [Fan, Meyer, and Simpson, 1960]. Thus, we may use the position of maximum intensity to give us the trace in the meridian plane of the geomagnetic line of force passing through E_3 (max) or E_2 (max).

THE GEOMAGNETIC FIELD AND THE REPRESENTATION OF EXPLORER VI ORBIT

The magnetic field of the earth is represented by a dipole inclined with respect to the rotation axis, plus higher order terms, and the period of the satellite is approximately 12.7 hr. As the satellite successively passes through different portions of the magnetic field in its orbit about the earth, only returning to its initial position with respect to geomagnetic coordinates after approximately 32 satellite orbits. For studies of the outer radiation belt the range R (distance from dipole to electron) is the order of 15×10^3 to 30×10^3 km. Therefore, we can neglect the higher order terms. The best test for the validity of this approximation is found in the consistency with the data obtained in the outer radiation belt with other smoothly with the dipole approximation over a wide range of geomagnetic coordinates at different times. This is shown as

Figure 8 we plot the positions of E_3 (max) in $\cos^2 \lambda$ coordinates for which the line of force of the center dipole approximation becomes a straight line passing through the origin. The data were obtained in a short time interval since

we find that the position of the intensity maximum changes with time. The high latitude data have been corrected according to Appendix 2. Within the limits of experimental errors, which we believe to be ± 300 km, the data lie on straight lines. The data for August 21 were obtained during magnetic disturbances (Fig. 13) when the intensity was changing by a large factor. Hence, we conclude that the centered dipole approximation is valid in the outer radiation belt not only for the undisturbed field but also during times of geomagnetic disturbances. This result has been shown in our earlier report [Fan, Meyer, and Simpson, 1960]. The important implication of the alignment of E_3 (max) with undistorted dipole lines of force is that the energy density ρ_e at all times must be less than the magnetic field energy density ρ_m in the outer radiation belt (for otherwise the magnetic field line would have been distorted by the presence of the particles).

Thus the problem reduces to the description of the satellite motion in the geomagnetic dipole meridian plane—a representation used also by other investigators in the description of satellite and space probe measurements. Figure 9 shows typical trajectories in the λ - R plane and indicates how, over an extended period of time, large volumes of space in the vicinity of the trapped radiation are covered by the Explorer VI orbit.

However, we wish to point out another representation for the orbit trajectory [Fan, Meyer, and Simpson, 1961]. If the dipole approximation is valid for outer-belt analysis as we have claimed, then the geomagnetic field lines are described by the equation $R = R_0 \cos^2 \lambda$, where λ is the geomagnetic latitude and R_0 the equatorial distance of a line of force from the center of the dipole. Lines of force with range R_0 around the earth are equivalent (azimuthal symmetry). Thus, the orbit may be described on a plane R_0 vs. R wherein each field line becomes a horizontal line in the two-dimensional plot. Trajectory data in these coordinates are shown in Figure 10. We see that a significant portion of the satellite motion may be projected onto a single dipole line of force whose range at the equator is R_0' . Since the electron intensity is measured continuously we may obtain curves for intensity as a function of λ (or magnetic field intensity) along a line of force as shown in

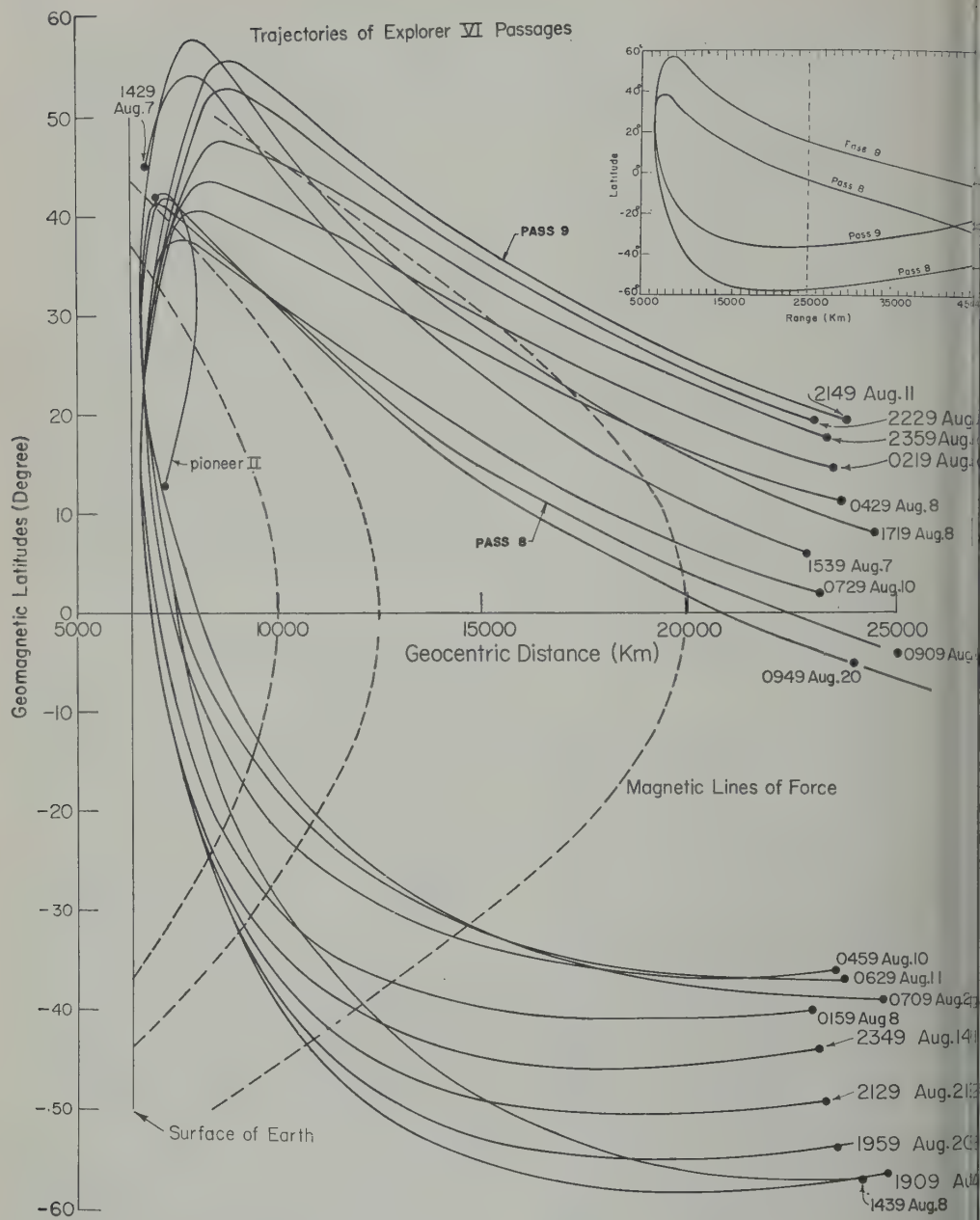


Fig. 9. Projection of satellite orbit on the meridian plane of the geomagnetic field using the centered dipole approximation. Insert shows the complete trajectory.

section 7. If we allow a range spread of about 200 km, we then obtain useful data of this type from over one-third of all Explorer VI trajectory passes. Among these passes we find

many that lie on the same R_0 as the position E_3 (max).

In section 7 it will be shown that the intensity along a line of force near E_3 (max) changes

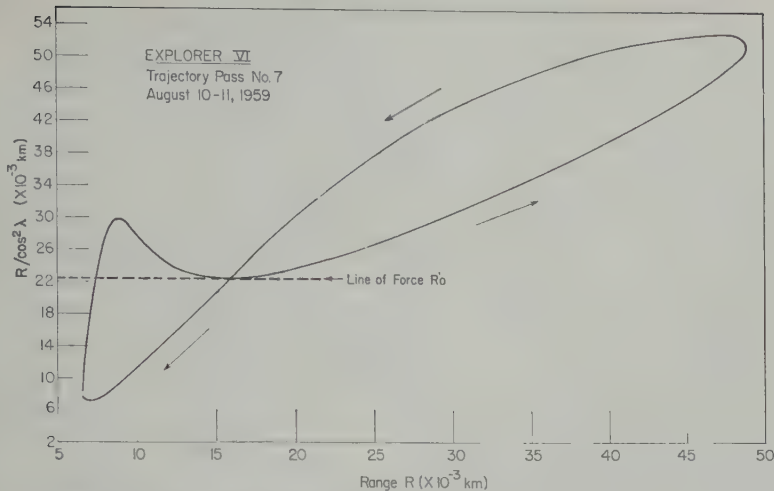


Fig. 10. Representation of Explorer VI trajectory in coordinates R_0 vs. R in the approximation of a centered dipole magnetic field where $R = R_0 \cos^2 \lambda$.

ly with λ . Also, we note from Figure 9 the satellite sometimes moves obliquely as the region of E_s (max) where the intensity is a strong function of latitude and range. In such cases the observed position of electron density peak and the true position of intensity maximum are in general different. This problem is discussed in Appendix B. The problem is restricted to high latitudes, since at low latitudes the traversals are always at a large angle with respect to the lines of force. From Appendix B it is clear that corrections in range and latitude are significant in locating the true position of E_s (max) at high latitudes. We have applied these corrections in the region of E_s (max) in this paper.

In Figure 11 the positions of the maxima for regions E_2 and E_s are plotted in the meridian plane for a large number of trajectory passes. This is an extension of results already published [Meyer, and Simpson, 1960]. From Figure 11 it is clear that the two maxima lie along lines of force and retain their separate identity to very high latitudes. It is also clear that the intensity maximum between E_2 and E_s in the outer belt is not the 'gap' described in the literature between the inner and outer belts.

THE RADIAL CHANGE OF ELECTRON BELTS

To investigate in greater detail the radial change of the two electron belts as a function of time, evident from Figure 11, we have used all

passes through E_2 and E_s with the criterion for identification of maxima outlined in Appendix B. We then determined the equivalent equatorial range R_0 of the magnetic dipole line of force on which each maximum was found to lie. The results of this extrapolation appear in Figure 12. Since systematic and progressive errors in the time-position data of the satellite trajectory could introduce 'drifts,' we have investigated the question of the accuracy of the orbital data, especially with respect to variable atmospheric drag at perigee. Although we know there are errors between the calculated and real positions of the satellite at small range—serious for inner belt studies—the errors in the outer region until at least September 9 are from 500 to 100 km in range (see Appendix B). Consequently, the large changes of range in Figure 12 represent real changes of the intensity maxima with time and lead to the conclusions:

1. There exist rapid and simultaneous changes in range of E_s (max) and E_2 (max) during geomagnetic storms. The magnitude of the change may be as great as 10 per cent of the total range from the dipole center.
2. Both regions tend to undergo the same inward direction of motion with the magnitude of change being much less for E_2 .

Whether outward motion occurs depends upon the interpretation placed on the data after the geomagnetic storm on September 3. Unfor-

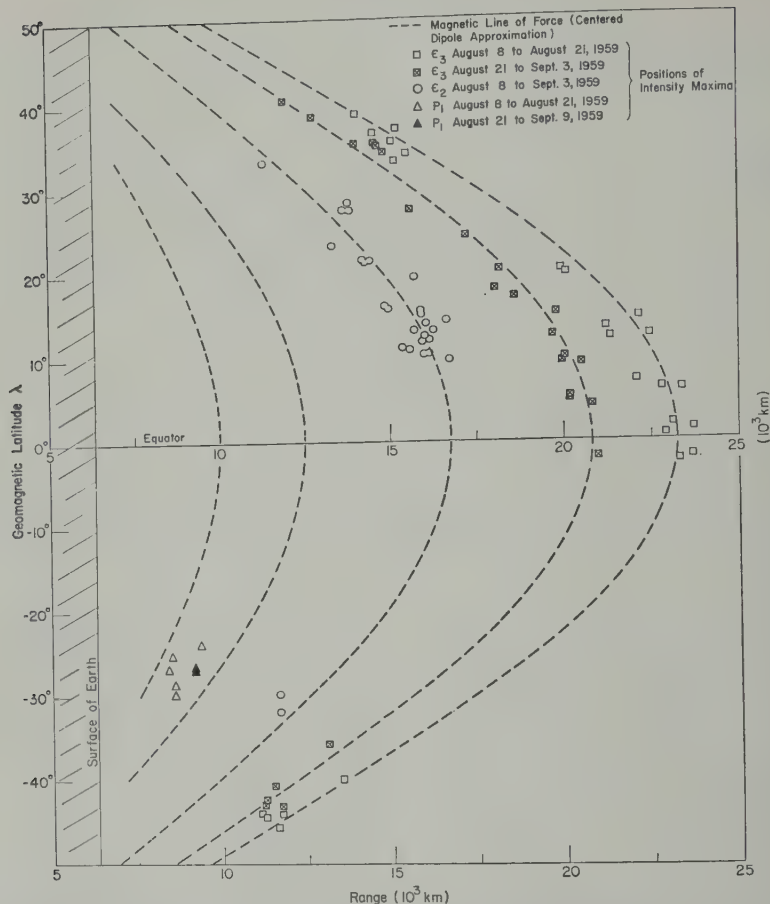


Fig. 11. Meridian plane plot of the E_2 (max) and E_3 (max) positions in the geomagnetic field. For more detailed analysis, see Figures 8 and 23.

tunately the ground stations were closed for a few days. We have indicated by arrows the alternative interpretations that may be placed on the data after September 3. Two regions continue to persist but their identities become obscure. The maximum appearing beyond 27,000 km after September 25 might be interpreted as either a new region forming and moving inward or a shift in the previous E_3 (max). In any event, inward motion is certain at some times, and outward motion is not excluded at other times by our results.

The physical explanation for these radial motions of the outer belt is unclear. If, as appears likely, the motions are mainly inward, then drift or diffusion across lines of force might seem to be an attractive explanation, especially since the drift occurred during a geomagnetic disturbance.

Although inward motion of electrons in uniform magnetic fields is a well-recognized possibility [Herlofson, 1960; Parker, 1961a], the short time constants and the persistence of a 'sharp' maximum with high intensity at times in the experimental data argue against inward diffusion.

The changes have the gross appearance of lines of force moving inward as though the strength of the dipole were decreasing. It has been argued that as a consequence of the insulating shell of the earth, lines of force may be progressively brought inward [Gold, 1959]. However, the distribution would rapidly disappear, contrary to fact.

The Argus experiment [Van Allen, McIlwain, and Ludwig, 1959b] showed that the intensity of maximum of trapped β rays from fission pro-

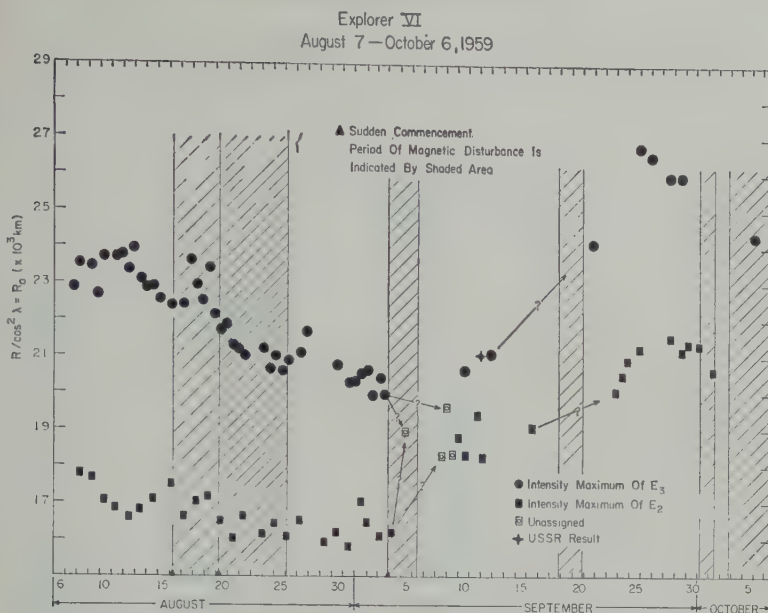


Fig. 12. The position of E_3 (max) extrapolated to the geomagnetic equatorial range R_0 as a function of time.

not change radially during the entire period of observation by more than ~ 30 km, even though geomagnetic disturbances were taking place. Since the ranges of the Argus 'shells' are less than for region E_2 , the radial motion we have found presumably has its origin in geomagnetic field and current system changes and the range of E_2 and E_3 at the times of geomagnetic storms.

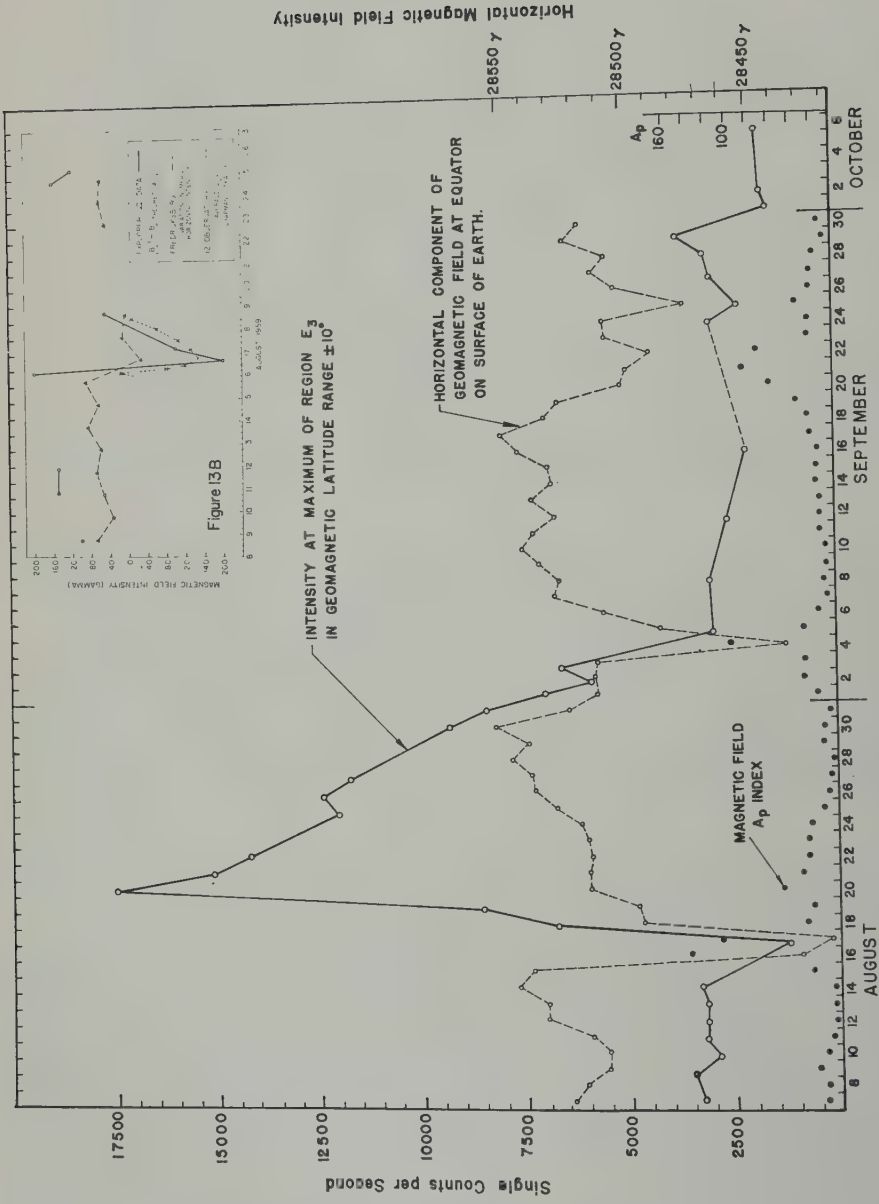
INTENSITY CHANGES OF E_2 AND E_3 REGIONS WITH TIME

Since satellite trajectories cover a wide range of geomagnetic latitudes as a function of time, the spatial distribution and changes with time of the intensities are intermixed. We shall first consider exclusively the data obtained in an equatorial band of latitudes $\pm 10^\circ$ for which the spatial variation may be neglected. We shall then extend this analysis to higher geomagnetic latitudes, and, finally, to a description of the intensity along a magnetic line of longitude as a function of time.

The results for intensity changes in the equatorial plane are shown in Figure 13 for the region E_3 . For the first 7 days after launch the intensity is constant within the accuracy of the measurements. Thereafter large intensity varia-

tions occurred in association with geomagnetic storms. On August 16, a severe geomagnetic storm began with sudden commencement at 0400 UT. The magnitude of the changes in the equatorial horizontal component of the geomagnetic field at the surface of the earth is plotted in Figure 13A. The magnetic field in the outer belt was also measured at this time (inset Fig. 13B) [Smith and Rosen, 1960; Smith and Sonett, 1961, to be published]. The period of violent magnetic disturbances, represented by large values of A_p , continued into August 17-18. Since it is well known that more than 7 to 10 days are required for the field to return to its normal condition, the second sudden commencement magnetic storm, which began at 0410 UT, August 20, must have been superposed on the recovery phase of the August 16 storm. The August 20 storm was coincident with a sudden decrease of ~ 15 per cent of cosmic-ray intensity in space [Fan, Meyer, and Simpson, 1960].

Gross changes of intensity in the E_3 region have also been observed with the other radiation detectors on Explorer VI [Arnoldy, Hoffman, and Winckler, 1960a, 1960b; Rosen, Farley, and Sonett, 1960]. The most obvious phenomena we note are a sudden decrease of intensity following the initial phase of the storm and, later, a



1959

Fig. 13A. The change of intensity of E_3 (max) in the geomagnetic equator ($\pm 10^\circ$) as a function of time. The averaged values of the surface horizontal component of the equatorial geomagnetic field (Huancayo, Peru) is shown. The A_p index indicating the magnitude of geomagnetic fluctuations is given as individual solid points. Changes in magnetic field intensity observed in Explorer VI within the outer belt [Smith and Rosen, 1960; Smith and Sonett,

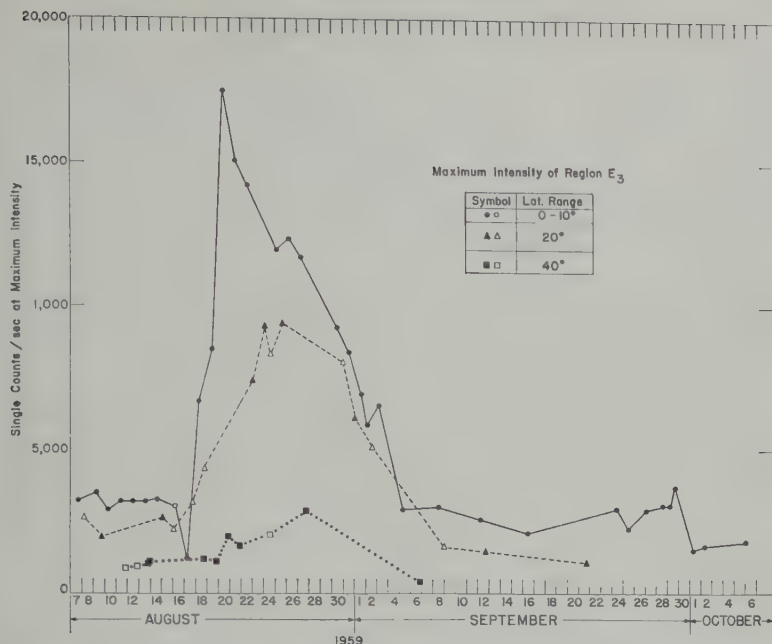


Fig. 14. Electron intensity as a function of time at E_3 (max) for different geomagnetic latitude intervals. The solid points represent data within $\pm 5^\circ$ of the 20° curve or the 40° curve.

dup by a factor 3 in intensity within 24 hours during the main phase and recovery of the first storm. The equatorial peak intensity decreased by another factor ~ 3 at the time of the second magnetic storm on August 20. The intensity then gradually declines, and, after the magnetic storm of September 3, reaches approximately the values obtained at the time of the first storm. We note here that geomagnetic storms occurred also on September 20 and October 3. These two storms were minor compared with the events of August 16 and 20.

We now investigate the intensity changes at time at higher geomagnetic latitudes, viz., at higher field intensities along lines of force at high E_3 (max). Since the intensity decreases rapidly with increasing latitude, curves similar to Figure 13 at higher latitudes can only be obtained by using data from small intervals of latitude. Therefore, all data in the latitude intervals $\lambda = 20^\circ \pm 10^\circ$ and $40^\circ \pm 10^\circ$ were interpolated (sec. 7) to 20° and 40° , respectively, and are shown in Figure 14. These curves for 20° , and 40° reflect the changes of electron pitch-angles with time. We discuss quantitatively the physical meaning of these observations after obtaining the detailed electron-intensity distribution along a line of force.

7. THE ELECTRON INTENSITY ALONG LINES OF FORCE AS A FUNCTION OF TIME

From the electron density measured along a tube of force, either the distribution of mirror points or the equatorial pitch-angle distribution of electrons may be determined. These measurements are fundamental to an understanding of the electron source and loss mechanisms in the outer belt.

We note from section 4, Figure 10, that for some passes the satellite moved along a dipole line of force for a significant interval of time, and that approximately one-third of all passes contained data of this kind. Curves for intensity I vs. λ have been constructed by supplementing these passes with individual crossings of the same force line at different latitudes but at closely related times.

We first show in Figure 15 the intensity along a line of force passing through the maximum of E_3 during times when the geomagnetic field is relatively quiet. Curve A is for a period before the geomagnetic storms, and curve B represents a quiet period³ where the counting rate was still high following the geomagnetic storms.

³ The lowest magnetic disturbance indices for August and September occurred on August 26.

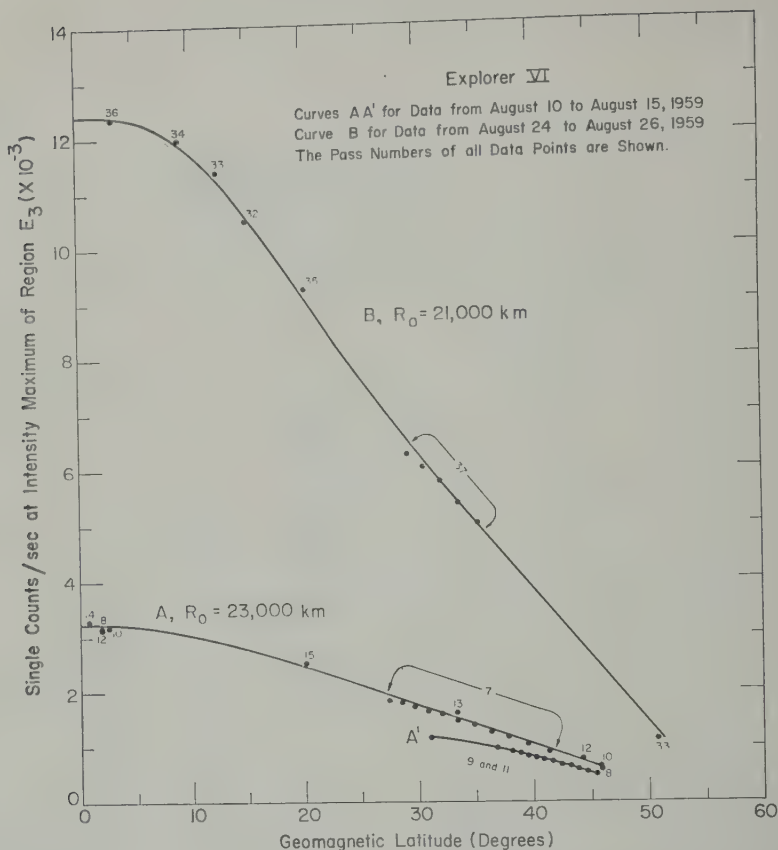


Fig. 15. Electron intensity at E_3 (max) along lines of force in the centered dipole approximation.

All additional data along segments of lines of force, not necessarily through E_3 (max), are given in Figure 16 for both magnetically quiet and geomagnetic storm periods.

We have found that all these data, to a good approximation, may be represented by

$$\frac{I}{I_0} = \left(\frac{B}{B_0}\right)^{-x} \quad (1)$$

at intermediate latitudes, and, for periods free from magnetic storms, even to the equator. The symbols I_0 and B_0 refer to equatorial electron and magnetic field intensity, respectively.

The data from Figures 15 and 16 have been plotted in Figures 17 and 18 using expression (1) and a dipole magnetic field for those cases where the lines of force passed through E_3 (max). The values of the exponent x for these curves are listed in Table 1.

Provided there are no large changes in the

electron spectrum with time along a line of force through E_3 (max), a condition we have shown to be satisfied (sec. 8, Fig. 21), then the curves in Figure 17 prove that the pitch-angle distribution (or mirror points) is changed during the main phase of the first geomagnetic storm. The relationship between the exponent x and the equatorial pitch-angle distribution has been discussed recently [Fan, Meyer, and Simpson, 1961]. Figure 17 also proves that the partial intensity increase is initially concentrated at the equator, and only at a later time in the development of the magnetic storm does the electron flux at large B/B_0 (high latitudes) begin to rise.

In the same way we investigate the region of the magnetic field where the lines of force are not closed. Unfortunately, owing to the high latitudes, the trajectory through the region of E_2 maximum is such that there are only individual points and no sections of data along lines of force. Therefore, we must be content with constructing the intensity a

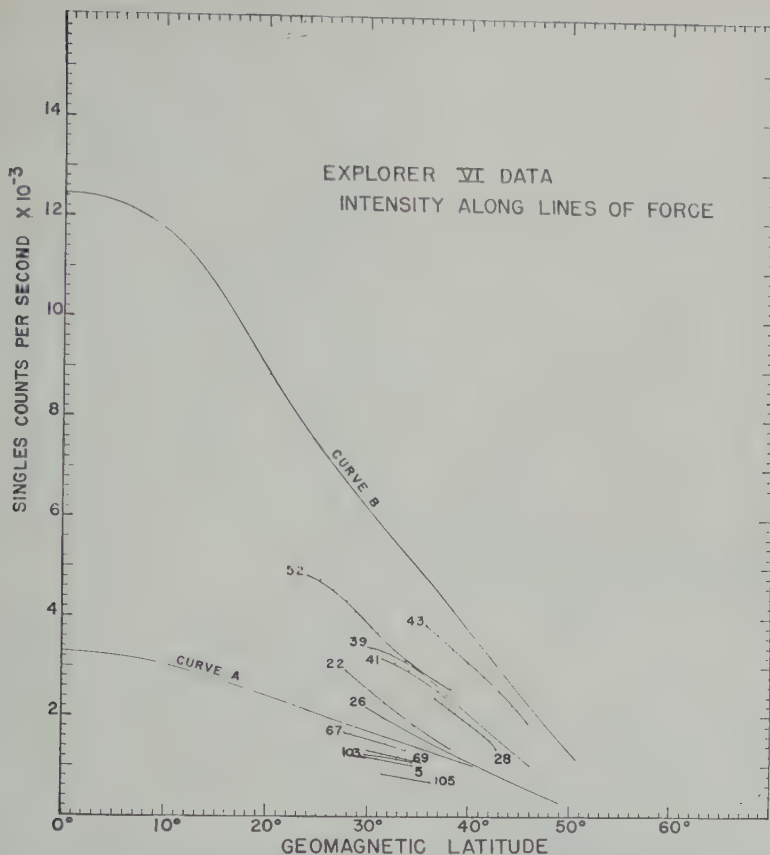


Fig. 16. All experimental data for E_3 (max) along lines of force similar to Figure 15, but covering a wide interval of time and range in space.

of force using data grouped over a period of several days. Since we showed [Fan, Meyer, and Simpson, 1960] that the region E_2 is relatively immune to the large-scale solar-induced phenomena associated with E_3 , we use data over intervals of time for which there was negligible change in range R_0 for E_2 . These time intervals— B_2 , C_2 , D_2 —are given in Figure 19. The assumption that the spectrum is independent of latitude no longer holds for region E_2 , as we show in section 8. Hence, we must confine our attention to the relatively large-scale changes in slope of the curves with time. The electron intensities along lines of force are shown in Figure 19 where the steep slope of curve B_2 corresponds to the time August 17–18 of the geomagnetic storm. We conclude that during this phase of the storm the mean pitch-angle of electrons in region E_2 also increased.

8. CHANGES IN ELECTRON ENERGY SPECTRUM WITH RANGE, LATITUDE AND TIME

The analysis of the outer-belt electrons has so far provided us with the time-dependent parameters of electron intensity I , equatorial range R_0 and E_3 (max), and a function x representing the electron pitch-angle, or mirror-point, distribution along a line of force. Since the electron flux is measured by its bremsstrahlung produced in the satellite, it is also essential to understand the major changes that may take place in the spectrum of trapped electrons during and after the magnetic storm. For example, if the spectrum of electrons changed along a line of force in such a way that the mean energy is much higher near the equator than at high latitudes during the magnetic storm (for example, during pass 22) it could not be argued

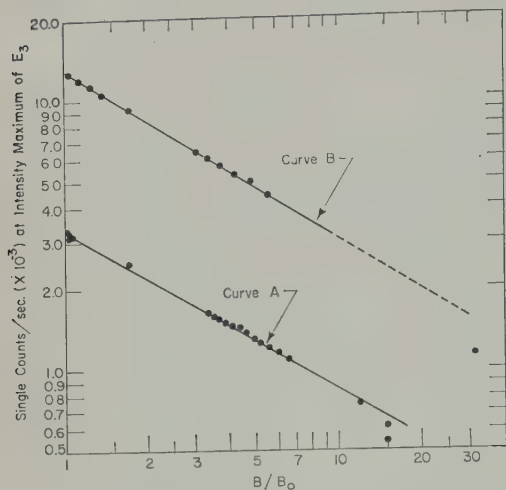


Fig. 17. Electron intensity along magnetic lines of force through E_3 (max) as a function of time. Using the data from Figure 15 and data during the geomagnetic storms August 16–20, it is shown that $I = I_0[B/B_0]^{-2}$.

that the observed increase in x (Fig. 17) was due to an increased concentration of particles at the geomagnetic equator.

We approach this problem of spectral change with time and position as follows. Although we do not know the electron spectrum we can determine if there are changes in the spectrum by observing the relative change in count rates of two bremsstrahlung detectors, each having a different response to the electron spectrum. Two detectors suitable for this purpose are the GM counter of the Minnesota Group [Arnoldy, Hoffman, and Winckler, 1960b] and the single, lead-shielded proportional counter.

Therefore, we use the change in the ratio

$$Z = \frac{\text{GM count rate}}{\text{proportional count rate}}$$

as an indicator of changes in the trapped electron spectrum.

The ratio Z is shown in Figure 20 as a function of equatorial range R_0 in 5° intervals of latitude. It decreases with increasing range R_0 but levels off at approximately 22,000 km. Thus, the electron spectrum at high latitudes and small range is quite different from the spectrum at E_3 (max). Beyond 22,000 km there is a negligible change in spectrum.

These conclusions are verified in Figure 21,

where we have plotted Z along selected lines of force. Pass 33 lies in the E_2 region. The data of pass 22 were obtained during the intense magnetic storm, thus proving that the intense changes along the line of force in Figure 17 are not due to changes in electron spectrum with latitude.

To isolate gross spectral changes with time from these spatial changes, we have selected only values of the ratio at E_3 (max) in the equatorial zone. This ratio is plotted in Figure 22 as a function of the single proportional counter intensity. We note here that (a) there are no significant changes in spectrum until after reaching maximum intensity at pass 25; (b) after pass 25 there is a systematic change of spectrum for many days (even though some of the variations in ratio may be due to changes in range etc.).

These results reveal that the electron spectrum for curves A and B in Figure 17 were different.

The above analysis did not depend upon assumptions regarding the form of the electron spectrum. If we assume, owing to lack of detailed information, that the electron spectrum is given by $N(E)dE = AE^{-\gamma} dE$, then we may predict specific changes in the energy spectrum. This has been done in Appendix C where the value of γ is derived from the count-rate ratios and the detector response curves.

9. THE ORIGIN OF CHANGES IN THE OUTER ELECTRON BELT

In the preceding sections of this paper we have established four parameters to represent the physical characteristics and changes with time of the outer electron belt, namely (1) the maximum electron intensity at the equator; (2) the mirror-point distribution along a line of force represented by the exponent x ; (3) the equatorial range of the dipole line of force passing through the principal maximum (E_3 (max)) of the outer belt R_0 ; and (4) the ratio Z which is sensitive to changes in the electron spectrum. All these parameters possess spatial and time-dependent characteristics.

We now investigate the basic question of the origin for changes in these parameters with time. It is clear that the time dependence originates with events at the sun which transport energy by means of plasma, radiation, and particles to induce geomagnetic storms. It

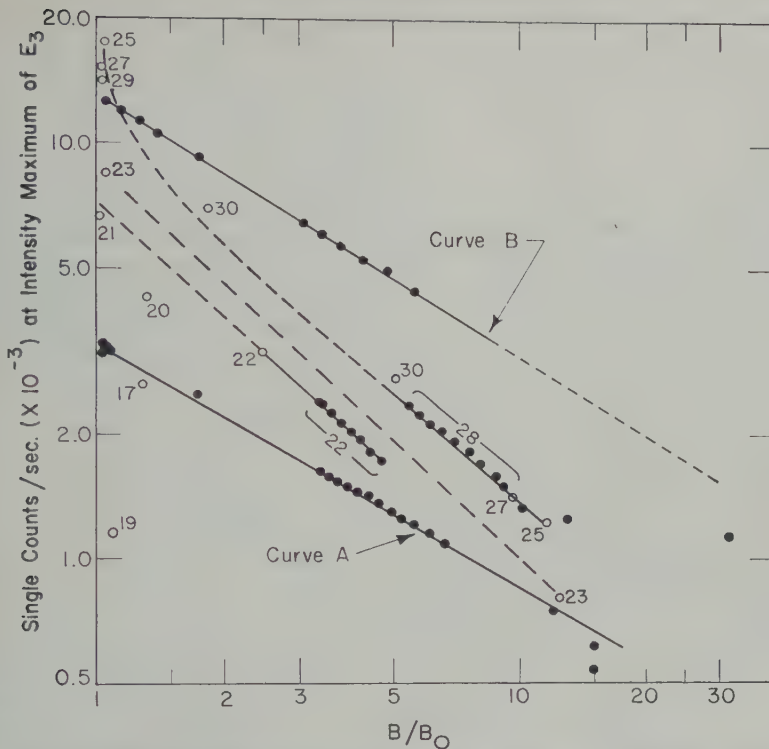


Fig. 18. Electron intensity along magnetic lines of force through E_3 (max) as a function of time later than August 26, 1959.

settled, however, whether the outer-belt electrons have a long trapping time and undergo changes in energy and trajectory through variations of the geomagnetic field in such a way that when the field returns to its initial condition, so will the parameters of the trapped electrons. Is the origin of the changes with time to be found entirely in reversible processes— for example, the betatron effect? Or, are we led to conclude that at least part of the change is brought about by particle loss, local electron acceleration or injection, and, there-

fore, are we required to admit that irreversible processes are prominent in the changing character of the outer electron radiation belt?

In this section we shall show how the analysis of the four parameters I_0 , R_0 , x , Z , along with a gross knowledge of the form of the time changes in geomagnetic field B , enable us to decide between these alternatives.

First, we bring together in Figure 23 the main results of our analysis in earlier sections for the period before, during, and after the two geomagnetic storms on August 16 and 20. This interval of time was chosen because of the large-scale effects, and the availability of large amounts of data.

Second, our analysis is restricted to the region E_3 where the position of the maximum intensity in the dipole field is used as a 'tracer' to follow the electron distribution with time. When we state that a particle distribution remains on a line of force through R_0 , we are using this line of force as representative of all lines of force around the earth at range R_0 across which these particles drift with time. We have assumed the

TABLE 1. Values of Exponent x as a Function of Time at Intermediate Latitudes

Pass, day	x	$R/\cos^2 \lambda$
Curve A (8-10 to 8-15)	0.58	23480
(8-18)	0.91	21480
(8-19)	0.95	23400
(8-21)	0.79	20820
Curve B (8-26)	0.63	20900
(8-29)	0.60	20790

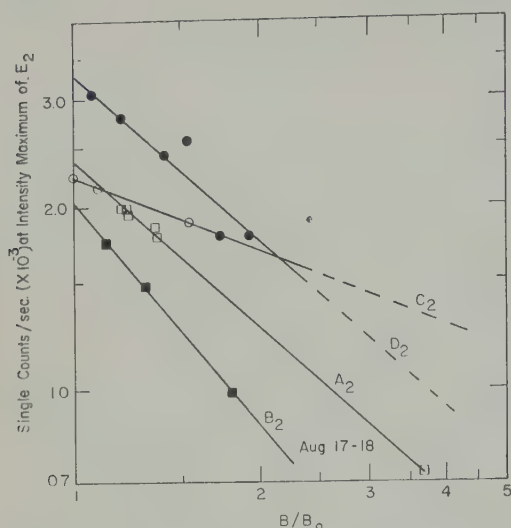


Fig. 19. For the region of E_2 (max) the intensity along lines of force have been determined by using data in the following time intervals: A_2 , August 7 to 15, 1959; B_2 , August 17 to 18; C_2 , September 8 to 11; D_2 , September 24 to 25.

centered dipole approximation where the field shape is independent of longitude.

Third, since the intensity changes of the geomagnetic field were measured in the E_3 region by the magnetometer on Explorer VI as shown in Figure 13B, we call attention here to the conclusion, important for the analysis that follows, that the magnetic field changes at the location of the E_3 region and at the surface of the earth (see Fig. 13A) are similar, especially with respect to the sign of the change in intensity with time. We also note that significant deviations of the magnetic field from the dipole approximation do not appear in the Explorer VI magnetometer records during the magnetic storms until ranges in excess of 25,000 km are reached [Sonett, Smith, Judge, and Coleman, 1960].

From Figures 13 and 23 we divide into eight successive time intervals the major changes in the outer radiation belt and relate them to the physical properties of the belt and the magnetic field.

Period t_1 (Aug. 7-15). This is the prestorm period where the mirror-point distribution was constant and the intensity changed by less than 1 per cent per day for the order of 7 days.

$$I_0(t_1) \approx \text{constant}$$

$$x(t_1) = 0.57$$

$$R_0(t_1) \approx 23,000 \text{ km}$$

$$Z(t_1) \approx \text{constant}$$

These represent the initial conditions for analysis.

Period t_2 (Aug. 16-17). The sudden commencement of a magnetic storm occurred at 0405 UT, August 16. Our first data after this time were obtained on pass 19 (August 16) during the main (depressed intensity) phase of the magnetic storm. The magnetic disturbance index A_p was high throughout this time period.

$$I_0(t_2) \approx \frac{I_0(t_1)}{3}$$

$$x(t_2) \text{ unknown}$$

$$R_0(t_2) \approx 22,000 \text{ km}$$

$$Z(t_2) \approx Z(t_1) \text{ (negligible spectrum change)}$$

We conclude that the changes between t_1 and t_2 could be explained by betatron deceleration.

Period t_3 (Aug. 17-18). The magnetic intensity at both the earth and the region of E_3 (max) had rapidly returned to within ~ 0.1 of its prestorm value for pass 21 on August 17. The magnetic disturbance index A_p was still high.

$$I_0(t_3) \approx 2I_0(t_1)$$

$$x = 1.2$$

$$R_0(t_3) \approx R_0(t_1)$$

$$Z(t_3) \approx Z(t_1)$$

Since (1) the range R_0 of E_3 (max) was the same as before the storm, (2) the magnetic field remained undistorted, and (3) the field intensity was still below the prestorm value, the increase of I_0 at t_3 by a factor > 2 above $I_0(t_1)$ can be explained by some irreversible process. The fact not only is $I_0(t_3) > I_0(t_1)$ but also $x(t_3) > x(t_1)$.

Period t_4 (Aug. 18-19). The magnetic intensity continues slowly to increase toward prestorm intensity on pass 23. The electron flux increases and a flux appears at great distances ($> 40,000$ km) as shown in section 10.

$$I_0(t_4) \approx 3I_0(t_1)$$

⁴ The STL scintillation detector count rate increased at this time. We discuss this increase in section 12.

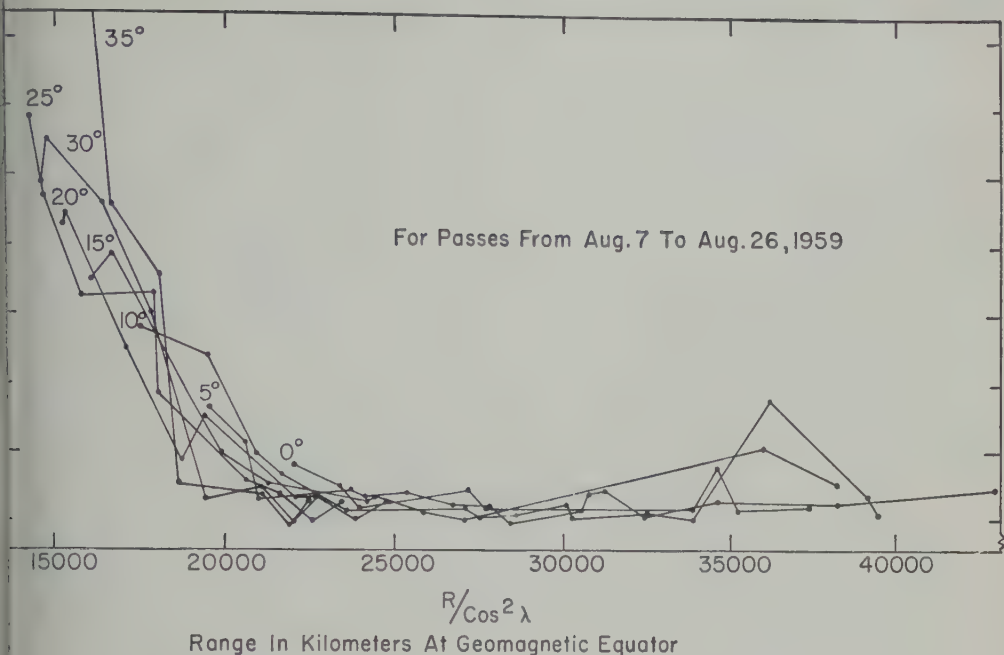


Fig. 20. The ratio Z of counting rates from two bremsstrahlung detectors as a function of equatorial range R_0 . All data points have been projected onto equatorial plane in the centered pole magnetic field approximation. A change in Z indicates a change in electron spectrum. LiM counter data are from Arnoldy, Hoffman, and Winckler [1960a, b] and private communications).

$$x \approx 1$$

$$R_0(t_4) = R_0(t_1)$$

$$Z(t_4) = Z(t_1)$$

step could be possibly accounted for by an effect.

At t_5 (Aug. 20). The second magnetic storm with sudden commencement at 0410 UT on Aug. 20 was followed within 4 hours by pass 25.

$$I_0(t_5) \approx 6I_0(t_1)$$

$$R_0(t_5) - R_0(t_1) \approx -2000 \text{ km}$$

$$x > 1$$

$$Z(t_5) \approx Z(t_1)$$

could account for the flux increase as a probable process if the field increase comes through the change in range of 2000 km (max). Note the high concentration of electrons confined to within $\pm 15^\circ$ of the geomagnetic equator at this time.

At t_6 (Aug. 21-26). Between t_5 and t_6

there are the following changes:

$$I_0(t_6) \approx 4I_0(t_1)$$

$$\approx 2/3 I_0(t_5)$$

$$x(t_6) = 0.6x \approx (t_1)$$

$$R_0(t_6) \approx R_0(t_5)$$

$$Z(t_6) > Z(t_1)$$

(This significant change of the spectrum⁵ is independent of latitude.)

The measurements for curve B centered in time on August 26 are for a period of magnetically quiet days even though the magnetic field intensity at E_s (max) was still below prestorm level. Note that the data suggest a lowering of mirror points and an increase in flux at high latitudes even though the field strength in this region continued to increase.

Period t_7 (Aug. 26-30).

⁵ See Appendix C for probable form of spectrum change.

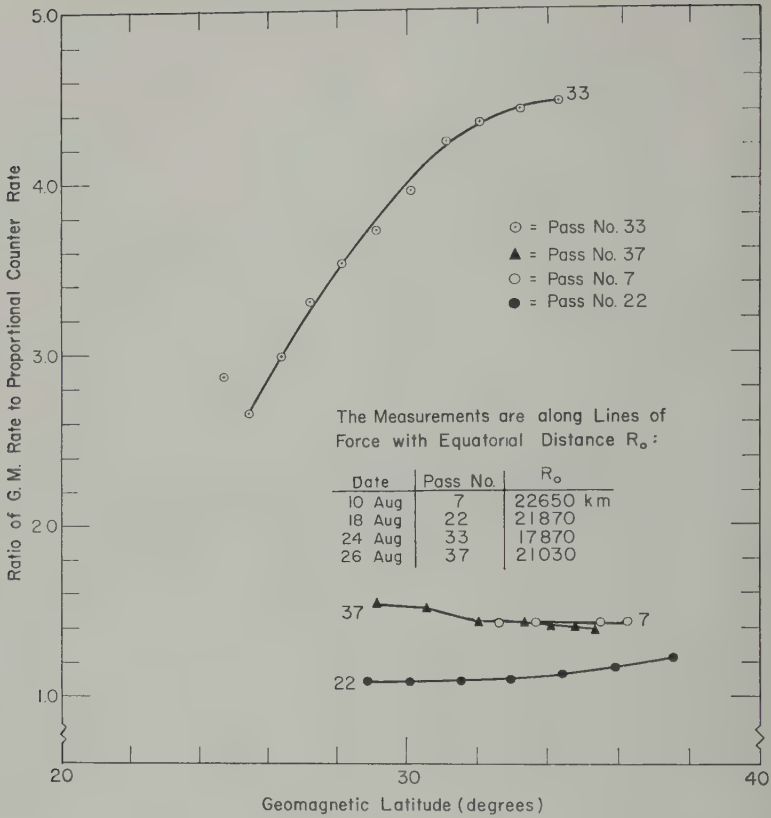


Fig. 21. The ratio Z along magnetic lines of force near E_3 (max) and E_2 (max).

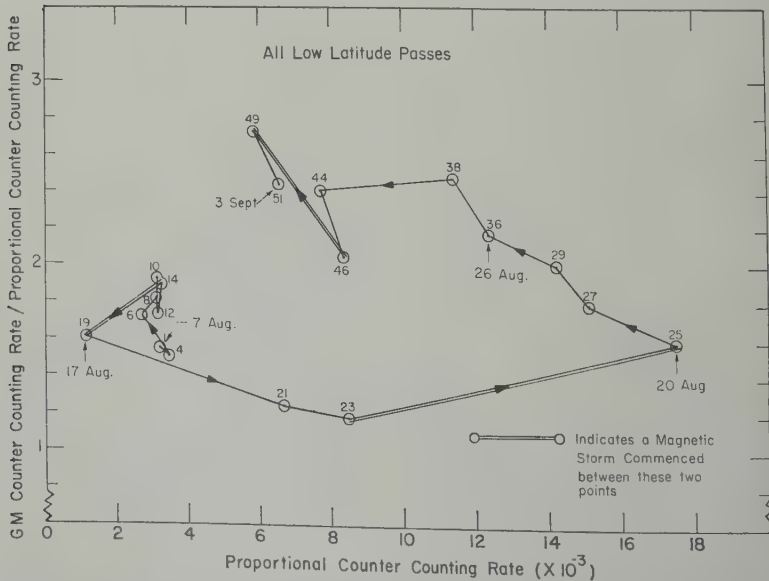
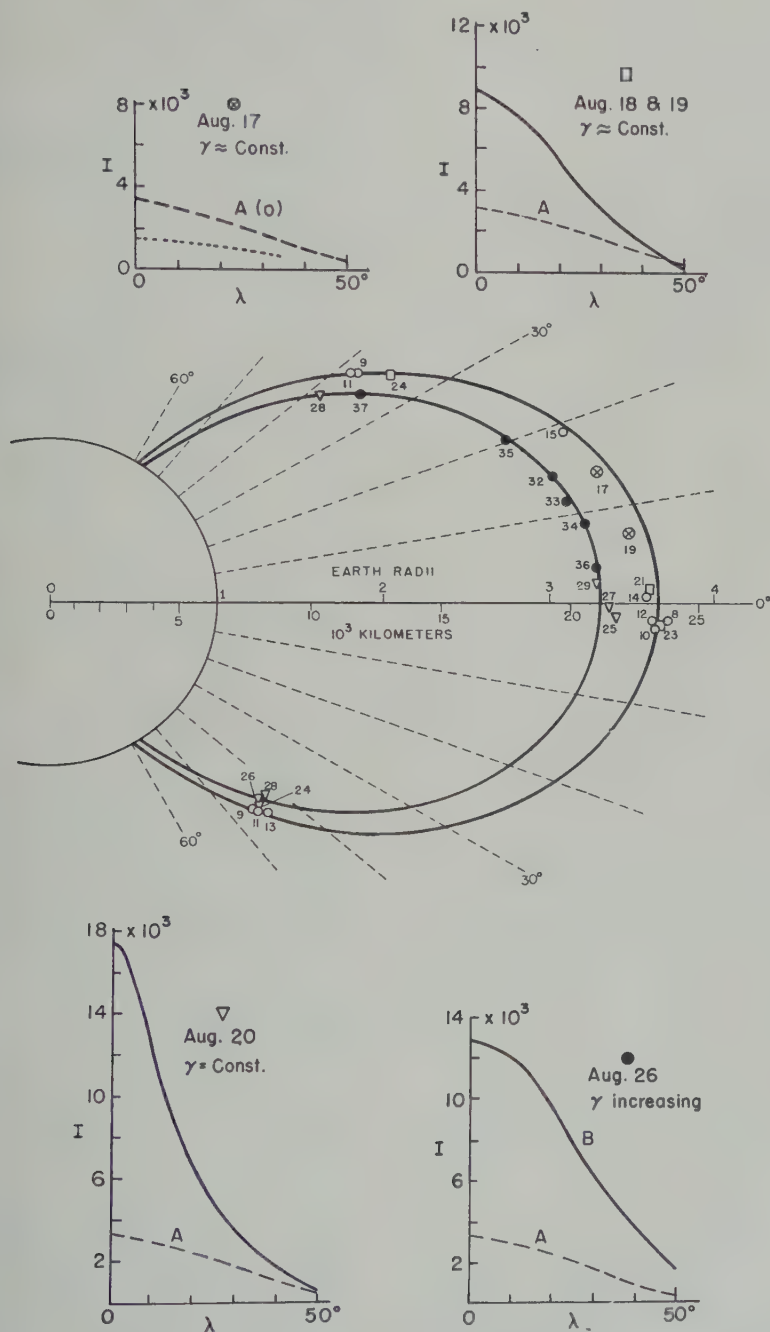


Fig. 22. The ratio Z as a function of the single proportional counter counting rate at E_3 (max).



These Data Are For The Maximum Intensity Of The E_3 Region

Fig. 23. The changes in pitch-angle (or mirror-point) distributions and intensity for electrons E_3 (max) as a function of range R_0 , and time for the geomagnetic storms of August 16 and 20, 1959.

$$I_0(t_7) = 2I_0(t_1)$$

$$R_0(t_7) \approx R_0(t_1)$$

$$x(t_7) = x(t_6) = x(t_1)$$

$$Z(t_7) \approx Z(t_6)$$

We find here a reduction of intensity by a factor of ~ 2 without a significant change in the pitch-angle distribution and the spectrum.

Period t_8 (Sept. 3). The magnetic storm of September 3 with sudden commencement at 2200 UT was followed by high A_p values and resulted in a rapid return of I_0 to the value $I_0(t_1)$ without a significant change in x as seen in Figure 16.

$$I_0(t_8) = I(t_1)$$

$$x(t_8) = x(t_1)$$

$$R_0(t_8) = \text{uncertain}$$

$$Z(t_8) \geq Z(t_7)$$

It is difficult to prove experimentally that a process is reversible because so much must be known about both the particle and magnetic field parameters. For example, *Parker* [1961b] recently has shown that the superposition of two deformations of the geomagnetic field can approximate almost any variation of electron distribution which maintains a 'bell-shaped' latitude dependence. Parker finds three adjustable parameters are required to describe the possible reversible changes in a slowly changing geomagnetic field: namely, $T(t)$, which determines the magnitude of the field B_s' , $a(t)$, giving the scale of the field; and $\nu(t)$ describing the form of the field lines. Accordingly, wherever we state that a given step could be explained as a reversible process, we imply the possibility of invoking changes in these magnetic field parameters to reconstruct the observed change in electron distribution.

The changes occurring between some of the eight periods above may be reversible. However, there are some changes which, under any model for trapped particles, could not be explained by reversible processes. Indeed, we need only prove that one of the series of changes is irreversible, and we prove the process as a whole to be an irreversible effect. For this test we choose first the change from period t_1 to period t_3 .

We note that between t_1 and t_3 the magnitude

of B decreased ($T(t)$ was decreased) while both the dipole character of the field lines and the range of the field line through E_s ($a(t)$ were preserved ($a(t)$ and $\nu(t)$ constant). Yet the electron density at t_3 was everywhere greater than at t_1 . This change in the radiation belt has its origin either in an increase in energy of already trapped electrons, or additional particles have been added in the energy range to which the detectors are sensitive.

What irreversible process could bring about this result? Briefly, there exist two external alternatives:

1. If, at period t_2 during the main phase of the geomagnetic storm with the total intensity at its minimum value, electrons were injected somehow with energies below the detection level of the bremsstrahlung detectors, then, when the field recovers to ~ 0.6 of its initial value, these new particles could be brought up to energies where they produce bremsstrahlung with increased efficiency. However, this explanation is unlikely since it would require to preserve the ratio Z (or the spectrum) and the new particles would have to be concentrated about E_s (max) to preserve the ratio of E_s (max).

2. During the period t_2 of the storm's main phase there were small-scale but rapid magnetic field intensity variations (high A_p) which could be invoked through some irreversible process to accelerate electrons. Such an unspecified process plus the partial recovery of over-all field intensity could not only bring the intensity above the prestorm value but also increase the electron density near to the equator.

We do not attempt here to build models for the observed irreversibility but only emphasize that the existence of this step removed the possibility for a reversible theory of outer magnetospheric intensity variations.

Another irreversible step clearly exists between t_5 and t_6 . On the basis of equatorial data the magnetic field intensity increases while the changing electron distribution remains almost on the dipole line of force of constant R_0 . The evidence is against any change in scale of the field. If the electron mirror-point distribution returns to nearly the prestorm values, which is equivalent to the equatorial intensity declining while the high latitude intensity increased. These changes were accompanied by a change in energy,

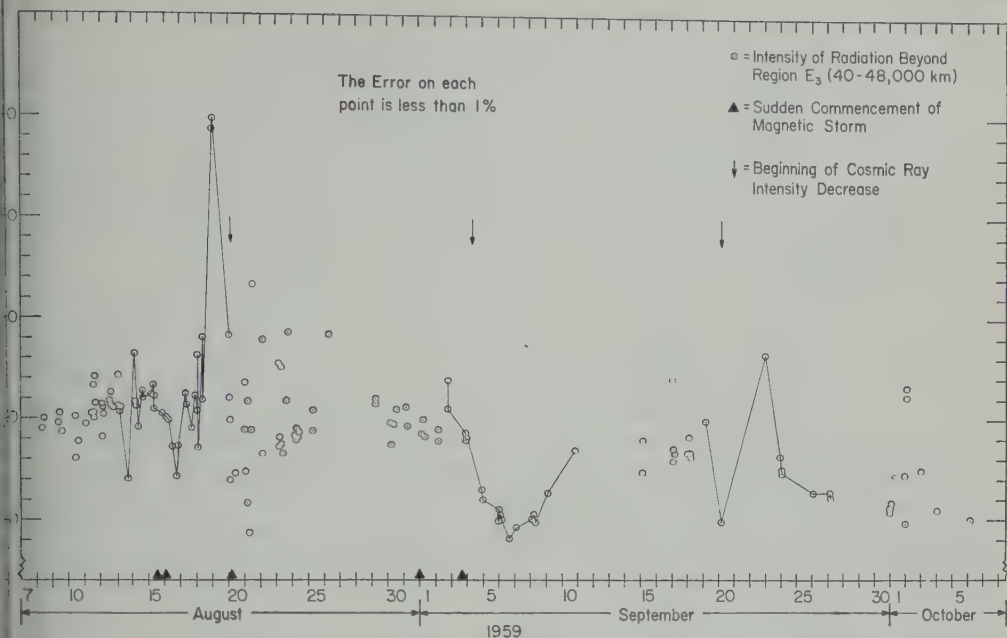


Fig. 24. The high-energy electron flux observed beyond the outer electron belt, i.e. at a range $R_0 \geq 40,000$ km as a function of time.

If we assume that the spectrum was of form $E^{-\gamma}$ then we observed an increase in γ . The changes reveal an irreversible loss of particle energy, or particles. The data indicate a net loss of particles at high energy which may come about either by the escape of higher-energy particles or a radiative energy loss at high energy. A similar irreversible process is observed between t_0 and t_1 .

From our investigation has led us to two irreversible processes, one an irreversible energy loss and the other an irreversible loss. At present we do not understand either of the underlying phenomena.

THE ELECTRON FLUXES BEYOND THE OUTER RADIATION BELT

At ranges greater than 40,000 km (corresponding to lines of force reaching the earth at magnetic latitudes greater than 67°) the scintillation counter normally detects only the cosmic-ray flux. Therefore, we may invoke a hypothetical, spherical surface with radius of 40,000 km to surround the trapped radiation and determine the fluxes of high-energy electrons, those that are trapped or pass through this 'boundary' from time to time. We do this by

plotting in Figure 24 the single count rate for all data at distances greater than 40,000 km as a function of time. As shown in section 2, this count rate is the sum of the omnidirectional cosmic-ray intensity, normally 9.5 counts/sec, and any bremsstrahlung produced by electrons. The singles count rate for the cosmic radiation exhibited three successive Forbush-type intensity decreases, as shown by the arrows in Figure 24. Thus, the excess count rate over approximately 10 counts/sec is due to bremsstrahlung by electrons incident on Explorer VI. One single count indicates an omnidirectional intensity of 3×10^6 electrons/sec/cm² if the energy is 200 keV. We note that in addition to low level fluctuations (less than 1 count/sec) there exist occasional outstanding bursts, especially the increase on August 19, which would represent a flux of 9×10^6 electrons/sec/cm² for 200 keV electrons.

The scintillation counter detected rapid intensity fluctuations correlated with magnetometer fluctuations in the same regions of space [Farley and Rosen, 1960]. These electrons must be trapped.

Since these electrons appeared during the period of intensity buildup in region E_3 , it is

tempting to assume that the flux represents the arrival of solar electrons, which diffuse inward in the equatorial magnetic field. However, they could not be preaccelerated electrons that find their way into the E_3 region for the purpose of producing the many-fold increase of equatorial intensity because (1) the E_3 intensity buildup began (between periods t_2 and t_3) many hours before this distant flux reached a high level and only after the kind of field fluctuations which might allow trapping had subsided; (2) there is no *a priori* mechanism whereby the process of trapping new particles should preserve the location of E_3 (max), or the energy spectrum; and (3) as a precondition for trapping, the incoming particle density must approximate the trapped particle density; but, on August 18–19 the flux was several orders of magnitude lower than required.

However, the possibilities exist that these particles are 'spillage' from the auroral zone, or that they have drifted outward from the equator as a result of instabilities⁶ during the rapid buildup of E_3 (max) intensity.

11. ELECTRON FLUX, ENERGY DENSITY, AND SPECTRUM

Bremsstrahlung measurements do not give directly the electron flux and energy spectrum. Therefore, we shall assume an energy spectrum and compute the electron flux from the single counter count rate and response curve in Figure 4 for monoenergetic electrons. For any assumed energy spectrum, we have the additional boundary condition that the energy density of the geomagnetic field ρ_m , at any point must considerably exceed the electron energy density ρ_e at that point. This condition was proved in section 4 for both normal and disturbed geomagnetic field conditions. This boundary condition, $\rho_m > \rho_e$, is most severe in the limiting case of the geomagnetic equator at the range of maximum E_3 intensity, where the electron density is highest and the field energy density is lowest along the line of force passing through E_3 (max). For the quiet period before August 14, 1959, where E_3 (max) lies at approximately 22 to 23×10^3 km, we find for an undisturbed

dipole field a magnetic field energy density $\approx 2 \times 10^{-8}$ erg/cm³.

Experimental evidence on the electron spectrum is scarce. The only direct measurement [Walt, Chase, Cladis, Imhof, and Knecht, 1961; Cladis, Chase, Imhof, and Knecht, 1961] carried out at low altitude (≈ 1000 km) a line of force passing through the geomagnetic equator at a range of ≈ 9000 km and, therefore, far below the range of E_2 (max).

In the energy range accessible to our detector their data are consistent with a differential energy spectrum of the form $N(E) dE = A E^{-\gamma} dE$ with $4 < \gamma < 5$. This exponent agrees with the analysis of the ratio of GM and proportional counters for a range less than the range of E_3 (max) as shown in Figures 20 and 25, Appendix. It then follows that if we continue to assume a power law spectrum to ranges beyond E_3 (max) then the exponent drops to $\gamma \approx 3$. Thus in the discussion which follows we shall assume a differential power law spectrum with $3 < \gamma < 5$.

We may determine the electron flux for the time interval August 7–14, 1959, at E_3 (max) in the equatorial plane as follows. The count rate, I , of a single proportional counter is given

$$I = \int_{E_1}^{E_2} N(E) G(E) dE$$

where $N(E)$ is the electron differential energy spectrum and $G(E)$ is the response function of our detector, shown in Figure 4. The lower limit E_1 is not critical for determining I since the function $G(E)$ falls off faster than E^{-3} below 100 kev. On the other hand, the value of E_2 is extremely sensitive to E_1 since the power spectrum is divergent as $E_1 \rightarrow 0$. For the present case we set $E_1 \approx 100$ kev. The upper limit E_2 is not well known. From Walt and his co-workers it should be less than 800 kev. From the results of Van Allen and the Soviet measurements it is consistent with a value less than 1000 kev. Therefore, we assume $E_2 = 1000$ kev. The parameter A in the expression $A E^{-\gamma}$ depends very little on the choice of E_1 , but is sensitive to the choice of E_2 . Increasing E_2 reduces A .

The values for the integral electron flux in the region E_3 at the geomagnetic equator are given for $\gamma = 3$ and $\gamma = 5$ in Table 2, along with the parameters used in the calculations. They show that the results are not very sensitive to the choice of the upper limit E_2 , flux value

⁶ For example, the recently studied velocity-space instability observed in mirror-machine geometries by R. F. Post (private communication to J. A. S.) [Post and Perkins, 1961].

TABLE 2. Parameters for Electron Flux and Energy Density in Region E_3 at the Geomagnetic Equator

γ	E_1 , keV	E_2 , keV	A $\frac{\text{electrons}}{\text{cm}^2 \text{ sec}}$ (keV) $^{\gamma-1}$	$N(E > 100 \text{ keV})$, electrons/cm $^2 \text{ sec}$	ρ_e , ergs/cm 3	ρ_m , ergs/cm 3
3	100	1000	1.6×10^{13}	8×10^8	4×10^{-8}	2×10^{-6}
3	100	2000	2.4×10^{12}	1.2×10^8	6.8×10^{-9}	
5	100	1000	5×10^{13}	1×10^{10}	8×10^{-7}	
5	100	2000	2.8×10^{13}	7×10^9	3.3×10^{-7}	2×10^{-6}
Equilibrium spectrum from neutron ray only*	0	800	5.3×10^{-5}	$N(E \geq 0 \text{ keV})$ 4.4×10^8	3.6×10^{-8}	2×10^{-6}

*Based on spectrum from Hess and Poirier (private communication to J. A. S.)

$E_1 = 1 \text{ Mev}$ and $E_2 = 2 \text{ Mev}$ are given. electron energy density is then,

$$\rho_e = \int_{E_1}^{E_2} \frac{N(E)}{v_e(E)} E dE$$

$$v_e(E) = c \frac{[\eta(\eta + 2)]^{1/2}}{\eta - 1}$$

velocity of the electrons, $\eta = E/m_e c^2$, m_e and c are the electron mass and the velocity of light, respectively. The computed values of ρ_e are given in Table 2.

There is some evidence [Vernov, Chudakov, Vernov, and Logachev, 1959] that there are variations in the spectrum below 100 keV, our calculations tend to be lower limits. We also note that the total intensity at E_3 (max) increased sixfold at the equator by August 21 under the condition $\rho_m > \rho_e$ (with the spectrum essentially unchanged as shown in Table 2). Thus it is clear that the exponent γ is much in excess of $\gamma = 3$, under the assumption of a power law spectrum.

In Table 3 a similar analysis in the region of E_2 (max) at $R_0 \approx 16$ –17,000 km shows that ρ_m even for the case of $\gamma = 5$.

12. SUMMARY AND DISCUSSION

Structure of the outer belt, the geomagnetic field, fluxes of electrons and protons. Measurements of bremsstrahlung in the satellite Explorer VI have revealed that for about 2 months there existed two distinct peaks of electron intensity, identified as regions E_2 and E_3 . This structure is a persistent feature of the

outer belt. It is proved that the maximum intensity in the E_3 region coincides over a wide range of geomagnetic latitudes with magnetic lines of force in the centered dipole approximation. Thus, the measured electron intensity maxima may be used as tracers of the geomagnetic field lines of force for the purpose of analyzing changes in the belt with time. During geomagnetic storms the electron intensity increased manyfold at E_3 (max), but even under these conditions the trace of the intensity maximum followed a center dipole line of force. This leads to the conclusion that the electron energy density must have been everywhere less than the magnetic energy density throughout the time of our measurements.

The observed range for the maximum electron intensity undergoes large and rapid inward shifts during geomagnetic storms which amount to as much as 10 per cent of the total range of the outer belt. This effect is likely to have its origin beyond the range of E_3 and connected with the irreversible energy gain in the outer belt discussed later in this section.

By introducing the assumption that the bremsstrahlung is produced by an electron spectrum of the form $E^{-\gamma}$, and by noting that the particle energy density is everywhere less than the magnetic energy density, we may calculate the maximum allowed electron fluxes in the region of E_3 under a variety of assumptions. For this assumed spectrum and energy limits of 100 to 1000 keV, we find that γ should be close to 3 and cannot be as large as 5. This agrees with independent calculations derived from the measured ratio of GM counter to proportional counter counting rate. Under these

TABLE 3. Parameters for Electron Flux and Energy Density in Region E_2

γ	E_1 , kev	E_2 , kev	$\frac{\Lambda}{\text{electrons}} \frac{(\text{kev})^{\gamma-1}}{\text{cm}^2 \text{ sec}}$	$N(E > 100 \text{ kev})$ electrons/cm ² sec	ρ_{e1} ergs/cm ³	ρ_{m1} ergs/cm ³
3	100	1000	1.1×10^{13}	5.3×10^8	2.8×10^{-8}	
3	100	2000	1.6×10^{12}	8×10^{-9}	4.4×10^{-9}	1 ×
5	100	1000	3.3×10^{13}	7×10^9	5.2×10^{-7}	
5	100	2000	1.9×10^{13}	4.7×10^9	2.2×10^{-7}	1 ×

assumptions the omnidirectional electron flux at $R_0 = 22,000$ km is 8×10^8 electrons/cm² sec. At smaller range in the E_2 region, a larger value for the exponent is allowed from energy density arguments, and is calculated from the observed change in counting rates of bremsstrahlung detectors. For the region E_2 a value of $\gamma = 5$ is consistent with high latitude observations.

The bremsstrahlung detector data show that the energy spectrum along a line of force E_3 (max) is essentially independent of latitude, but through E_2 (max) is sensitive to geomagnetic latitude, being strongly energy dependent at high latitudes. For example, if the electron energy spectrum has the form $E^{-\gamma}$ with an upper energy limit between 1000 and 2000 kev, then the form of the spectrum in the outer belt tends to become flatter (decreasing γ) with increasing range in the outer belt. The electron spectrum near E_3 (max) was observed to remain essentially unchanged even though the intensity changed over a factor of 6 as a result of a geomagnetic storm. At a later time, the spectrum was observed to change significantly and in such a way that for our assumed spectrum the value of the exponent increased by 25 to 50 per cent within a few days, and persisted long after the termination of the magnetic storm. Thus, it is evident that geomagnetic storms may induce changes of the outer-belt electron spectrum as well as changes in total intensity.

The proton flux in the outer belt for energies in excess of 75 Mev is (0.0 ± 0.1) protons/cm² sec.

The origin of the outer-belt electrons. We have described properties of the outer electron belt (sec. 9) by four parameters representing equatorial intensity $I_0(t)$, location of the magnetic dipole line of force through E_3 (max) at the equator, R_0 , the electron density distribution along a line of force $I = I_0(B/B_0)^{-\alpha}$ charac-

terized by the parameter α , and a measure of spectral change of electrons with time. We note that within the sequences of changes in the outer belt induced by geomagnetic storms, there are some changes of these parameters which cannot be accounted for only by involving an irreversible energy gain or energy loss within the belt. Since we measure bremsstrahlung, our rate I (single counter) is given by

$$I = \int_0^\infty G(E) N(E) dE$$

where $N(E)$ is the spectrum and $G(E)$ is the detector efficiency. Although we know $G(E)$ (see Fig. 4), we do not know $N(E)$, and therefore cannot prove whether a change in I arises from the gain or loss of particles, or whether it arises from a kinetic energy gain or loss for a constant number of particles. Thus, we confine our conclusions to the proof that there exist two kinds of irreversible processes: namely, (1) particle or energy gain to the belt; (2) particle or energy loss, leading to a stable intensity and electron mirror-point distribution.

Regarding the process of energy gain we have found arguments that strongly favor the acceleration of the electrons. Some of the arguments are of the negative kind. The mere existence of the high-energy electron flux beyond the outer belt proves that preaccelerated electrons are not present with sufficient density even if it were possible to trap them efficiently. An alternative mechanism of electrons at subdetection energies (100 kev) followed by local acceleration is unlikely because the electron distribution of the outer belt particles would have to maintain the constant E_3 (max) location in the magnetic field and the resultant energy spectrum unchanged.

We are led, therefore, to consider mainly the acceleration of electrons already trapped

here and at what times during the magnetic storm could this irreversible energy gain occur. The time of occurrence is fairly well limited to the first part of the main phase and extending through the first hours of recovery from the main phase of the geomagnetic storm (August 19, as shown in Figs. 13 and 23).

The mechanism for energy gain in this period is not clear. For example, if a statistical acceleration process, such as Fermi acceleration, were operative at this time, while small-scale magnetic disturbances were moving through the dipole region, we might expect that protons would be more efficiently accelerated than electrons. It is not likely that the sudden increase of electron intensity observed in the Explorer VI scintillator [Rosen and Farley, 1961] could be due to new energy protons beginning at the time of the all bremsstrahlung detectors and the ionospheric still observed low intensity on August 19. The scintillator detects protons of energy greater than 10 Mev. It may also be argued that the energy gain is coincident with the rapid recovery of the geomagnetic field—such as ring current intensity—on August 17–18. In any event, any successful model of acceleration must include (1) preservation of E_3 (max) over an intensity increase of factor 6; (2) no drastic changes in the shape of the spectrum; and (3) a concentration of the excess intensity in the vicinity of the minimum energy.

The loss processes are of a different kind. They are associated with a changing spectrum, a decrease of electron pitch angles and a gradual decrease of intensity. All these changes occur while the geomagnetic field gradually increases to the prestorm level and becomes unperturbed. The loss process seems to go on with a longer time constant than the irreversible acceleration described above. Again we are not sure whether the electrons are escaping from the trapped distribution or whether we are observing particle acceleration. The changes are consistent with the latter alternative, especially if we assume a loss of particles of the form E^{-1} . We then interpret the changes of particles or energy in two successive stages: first, there is a loss at high pitch angles and high energy, and second, there is a loss at low altitudes independent of pitch angle.

The changes take place as though there were some instability, or energy loss process (e.g., synchrotron radiative) which returns the electron

intensity distribution to the 'steady state.' In fact, one of the most remarkable results of our analysis is the observed tendency for the E_3 region to stabilize around a characteristic pitch-angle distribution where $x \approx 0.6$ except during magnetic field perturbations. Note that $x \approx 0.6$ even for the period after the September 3 magnetic storm (Fig. 16).

The fluctuations of high-energy electron intensity beyond the outer belt introduce the possibility that there may also be a loss of particles in near-equatorial orbits on or about August 20 when the equatorial concentration of electrons was greatest.

It has been suggested [Arnoldy, Hoffman, and Winckler, 1960b; O'Brien, Van Allen, Roach and Gartlein, 1960] that the large reduction in observed bremsstrahlung intensity with the main phase of a magnetic storm (in our case, pass 19) is in reality 'dumping' of particles to form the aurora. We believe that the small energy density of high energy electrons cannot account for the much larger auroral energy release, and that betatron deceleration between t_1 and t_2 is more likely.

The results are consistent with the acceleration of ambient electrons to the observed energy distributions. Probably through collective motions of high-temperature solar plasma somewhat more energetic particles could be trapped from time to time, but their equilibrium energies and distribution would still be dominated by processes in the geomagnetic field. In view of our conclusions, the question of the origin of energy gain and loss processes is more significant than the question: from where do the electrons come?

The beta decay electrons from neutrons certainly contribute to the outer belt [Hess, 1960; Dessler and Karplus, 1960]. Recently, however, it was shown that processes other than injection of neutron decay electrons must be operative in order to account for the spatial distribution of electrons in the outer belt [Hess, Killeen, Fan, Meyer, and Simpson, 1961]. This again points to the existence of energy gain and loss processes.

It is important to note that satellite bremsstrahlung observations in low altitude orbits, such as Explorer VII, will appear to have a time dependence different from our equatorial results for the outer belt. This is readily understood from Figures 14 and 23 for the August

1959 magnetic storms. For example, the increase of intensity would appear at the 'tip' of the outer belt a few days after reaching maximum intensity at the equator. The position in latitude of the maximum intensity near the 'tip' should shift rapidly by $\sim 1^\circ$ to 3° corresponding in our analysis of changes in R_0 (note that for low altitudes the center dipole approximation is not sufficient to describe the real field).

Finally, we caution the reader to recall that we have been able to analyze only two magnetic storms. We would expect the detailed changes in the outer electron belt from one magnetic storm to another to be as diverse as geomagnetic storms themselves have proven to be through many years of analysis by many investigators. Indeed, the magnetic storm of September 3 (for which we have very few data) did not lead to an irreversible acceleration, but only a loss mechanism which preserved the prevailing electron pitch-angle distribution.

Acknowledgements. It is a pleasure to acknowledge the essential contribution to the engineering of our apparatus in Explorer VI by L. Petraitis, R. Takaki, J. Jewzeski, W. Six, and J. Lampert of the Universities Laboratories for Applied Sciences. For data reduction and analysis we thank M. Case, C. Gloeckler, K. Schwarz, R. Langford, and G. Lentz. A continuous watch for unusual solar events by R. Tjonaman and the staff of the Cosmic Ray Group assisted in increasing the data acquisition during August 1959. The assistance of R. D'Arcy in the X-ray calibration of the duplicate payload was appreciated. We thank R. S. Rochlin and the General Electric Company for the use of their linear accelerator, and D. Yovanovitch for assistance in carrying out the proton calibrations.

Discussions of our results with E. N. Parker have been both stimulating and helpful to us.

We appreciate the collaboration among those conducting experiments in Explorer VI, especially the Minnesota group of Arnoldy, Hoffman, and Winckler and the Space Technology group of A. Rosen, T. Farley, and E. Smith, where we have mutually exchanged data. We thank the Minnesota group for permission to use their unpublished GM counter data and for the use of their X-ray apparatus.

In carrying out the Explorer VI project we are especially indebted to J. Lindsay and H. Newell of the National Aeronautics and Space Administration, and to A. Thiel, J. Lindner, and K. Moe of Space Technology Laboratories.

This research was supported in part by the National Aeronautics and Space Administration under contract NASw-24, and in part by the Air Force Office of Scientific Research under contract AF-18(600)-666.

APPENDIX A

TABLE 1. Characteristics of Charged-Particle Detectors and Magnetometer on Explorer VI

Proportional Counter Telescope	University of Chicago
A. Triple coincidence counts measure protons ($E > 75$ Mev) and electrons ($E > 13$ Mev) and are not affected by intense bremsstrahlung.	
B. Single counts measure bremsstrahlung and particles under A.	
GM Counter and Ion Chamber	University of Minnesota
A. GM counter measures sum of protons ($E > 36$ Mev), electrons ($E > 2.8$ Mev), and ambient bremsstrahlung.	
B. Ion chamber measures average ionization per particle (protons $E > 24$ Mev, electrons $E > 1.6$ Mev) plus bremsstrahlung.	
Plastic Scintillation Counter	Space Technology Laboratories
A. Measures electron flux directly for $E > 200$ Kev; plus	
B. Bremsstrahlung; plus	
C. Protons with $E > 2$ Mev	
Spin Magnetometer	Space Technology Laboratories
Measures component of magnetic field perpendicular to the spin axis of Explorer VI down to 10^{-4} gauss.	

TABLE 2. Low-Latitude Traversals Through the Outer Radiation Belt Which Yielded Data

	Dates	Number of Passes Through Outer Belt
1. Pioneer I	Oct. 11, 1958	
2. Pioneer III	Dec. 6, 1958	
3. Lunik I	Jan. 2, 1959	
4. Pioneer IV	Mar. 3, 1959	
5. Explorer VI	Aug. 7-Oct. 6, 1959	113
6. Lunik II	Sept. 12, 1959	
7. Pioneer V	Mar. 11, 1960	

BLE 3. Times for Orbit Pass Numbers of Explorer VI (Aug. 7-Oct. 6, 1959)

Pass No.	Date	Pass No.	Date	Pass No.
1	8-28	41	9 19	81
2	29	42	19	82
3	29	43	20	83
4	30	44	20	84
5	30	45	21	85
6	8-31	46	21	86
7	9- 1	47	22	87
8	1	48	22	88
9	2	49	23	89
10	2	50	23	90
11	3	51	24	91
12	3	52	24	92
13	4	53	25	93
14	4	54	25	94
15	5	55	26	95
16	5	56	26	96
17	6	57	27	97
18	6	58	28	98
19	7	59	28	99
20	7	60	29	100
21	8	61	29	101
22	8	62	30	102
23	9	63	9-30	103
24	10	64	10- 1	104
25	10	65	1	105
26	11	66	2	106
27	11	67	2	107
28	12	68	3	108
29	12	69	3	109
30	13	70	4	110
31	13	71	4	111
32	14	72	5	112
33	14	73	5	113
34	15	74		
35	15	75		
36	16	76		
37	16	77		
38	17	78		
39	17	79		
40	18	80		

APPENDIX B

process of determining the position of a radiation belt. We define the position of a radiation belt as the location in the geomagnetic equatorial plane at which the counting rate of a radiation belt is a maximum.

Consider a line of force of the geomagnetic field. A position on this line may be specified by the value of a parameter $b = B/B_0$ where B is the magnetic field intensity at this position, while B_0 is the minimum magnetic field intensity

along the line of force. Since an equatorial surface is a surface generated by the points of minimum field intensity on lines of force, i.e., the B_0 surface, the constant B_0 lines and their orthogonal lines, denoted by constant ϕ , may then be conveniently used as a coordinate system to specify the positions of lines of force. Thus, the three variables, B_0 , ϕ , and b determine a point in space, and we can write the counting rate as

$$I = I(B_0, \phi, b) \quad (1)$$

The variation of counting rate per unit arc length ds along the trajectory of the satellite is

$$\frac{dI}{ds} = \left[\frac{\partial I}{\partial B_0} \right]_{\phi, b} \frac{dB_0}{ds} + \left[\frac{\partial I}{\partial \phi} \right]_{B_0, b} \frac{d\phi}{ds} + \left[\frac{\partial I}{\partial b} \right]_{B_0, \phi} \frac{db}{ds} \quad (2)$$

If there is an axial symmetry so that $(\partial I / \partial \phi)_{B_0, b} = 0$, equation 2 becomes

$$\frac{dI}{ds} = \left[\frac{\partial I}{\partial B_0} \right]_{\phi, b} \frac{dB_0}{ds} + \left[\frac{\partial I}{\partial b} \right]_{B_0, \phi} \frac{db}{ds} \quad (3)$$

The physical meanings of $[\partial I / \partial b]_{B_0, \phi}$ and $[\partial I / \partial B_0]_{\phi, b}$ are respectively the variation of the counting rate along a line of force and that along the intersection of a constant b and a constant ϕ surface. The positions of the maxima shown in Figures 5, 6, 7a, and 7b are the solutions of $dI/ds = 0$ and are called the observed maxima, whereas the true maximum of the radiation belt by our definition is given by the equation

$$\left[\frac{\partial I}{\partial B_0} \right]_{\phi, b} = 0.$$

This is, of course, correct only if the pitch-angle distribution, $n(v, \theta_0)$ [Fan, Meyer, and Simpson, 1961] does not vary greatly with respect to B_0 .

The following is a working process by which the solution of $(\partial I / \partial B_0)_{\phi, b} = 0$ may be found.

Figure 17 shows that for a limited segment along a line of force of a center dipole field, equation 1 assumes the following simple expression

$$I = f(R_0)(b)^{-\lambda} \quad (4)$$

where, if R and λ are respectively the range and the magnetic latitude,

$$R_0 = R / \cos^2 \lambda, \quad b = \frac{(1 + 3 \sin^2 \lambda)^{1/2}}{\cos^6 \lambda} \quad (5)$$

and x is in general a function of B_0 . B_0 is then M/R_0^3 . Therefore, if the magnetic field in the vicinity of the *short portion of the trajectory* is not greatly different from a center dipole field, and if the value of x is almost constant, the position of the maximum of $f(R_0)$ ($=Ib^x$) will be approximately the solution of $(\partial I/\partial B_0)_{\phi, b} = 0$.

In the determination of the maximum of $f(R_0)$ we encountered two difficulties: (1) the determination of the position of E_3 required a knowledge of the value of x for $R_0 \approx 22,000$ km, which is not always available; (2) the position data errors are determined by the accuracy of the trajectory data discussed below.

The values of x during the quiet period of August 7-14, 1959, at $R_0 = 22,000$ km were 0.5 to 0.6, whereas in the magnetically disturbed days (August 16-20) they increased to about 1.0-1.2. This is about the total range in the measured values of x . The following criterion was used for the actual determination. If the value of x for certain passages at R_0 between 19,000 and 24,000 km was available, this value was used; otherwise, the value of x of the passages immediately preceding or following was used, provided that no magnetic storms were under way.

Take passage 20 as an example. The maximum counting rate along the trajectory ($dI/ds = 0$) is at $\lambda = 12.4^\circ$ and $R_0 = 24,500$ km, whereas the solution of $df/dR_0 = 0$ is at $\lambda = 14.4^\circ$ and $R_0 = 23,600$ km.

We have used improved trajectory calculations prepared by the staff at Space Technology Laboratories which take into account the variable atmospheric drag at perigee. These data appear to be accurate within $\approx \pm 100$ km at the outer belt until at least September 12. However, a new trajectory for Explorer VI will be prepared based on minitrack data providing more accurate position-time data at small range. In view of these trajectory errors the corrected positions of E_2 and E_3 have been worked out only for the cases where the solutions of $dI/ds = 0$ and $df/dR_0 = 0$ are widely different. No corrections are needed for values of $\lambda \leq 10^\circ$, since $B/B_0 \approx 1$.

Two final remarks must be added:

1. Since

$$\frac{d}{ds} \left[I \left(\frac{B}{B_0} \right)^x \right] = \frac{df}{dR_0} \frac{dR_0}{ds}$$

there always exists an extreme value $(B/B_0)^x$ at the position where $dR_0/ds = 0$; this is in general different from the solution of $df/dR_0 = 0$.

2. It can readily be shown that the 'ridge' equal counting-rate contours, frequently used as the definition of the position of a radiation belt, coincides with a line of force passing through the solution of $df/dR_0 = 0$. Therefore these two definitions are equivalent.

APPENDIX C

Derivation of changes in electron spectrum in bremsstrahlung detectors. We showed in section 11 that whatever may be the form of the electron spectrum in the outer belt the spectrum *uniquely* goes *changes* with range and along lines of force in the vicinity of the E_2 region. However, the spectrum was approximately constant along lines of force through E_3 (max).

We shall now introduce an assumed form for the spectrum and calculate the changes in the ratio Z of count rates from two bremsstrahlung detectors, each having different responses to a given energy spectrum. The response curves of the single proportional counter has been described in section 2 and Figure 4. The response curve of the GM counter in Explorer VI has been made available to us by Arnoldy, Hoffman, and Winckler. (The three groups of investigators using Explorer VI for radiation measurements have exchanged their calibration and detector characteristics.)

From section 11 it appears that a reasonable form of the electron energy spectrum prevalent in the outer belt is $dN(E) = AE^{-\gamma} dE$, where we shall use in our calculations. Our second assumption is that the maximum energy for outer belt electrons is $E < 2.8$ Mev, i.e., direct electron detection is possible by electron detector (Arnoldy, Hoffman, and Winckler, private communication).

We may then calculate directly the expected values of γ for given values of the ratio Z . The solid curve shown in Figure 25 uses the measured response curve in Figure 4 and a maximum electron kinetic energy of 1000 kev. To demonstrate the sensitivity of these calculations to the shape of the response curve we have also calculated Z vs. γ for the case where the counter response is proportional to $E^{1.7}$ over the entire energy range. Clearly, the changes

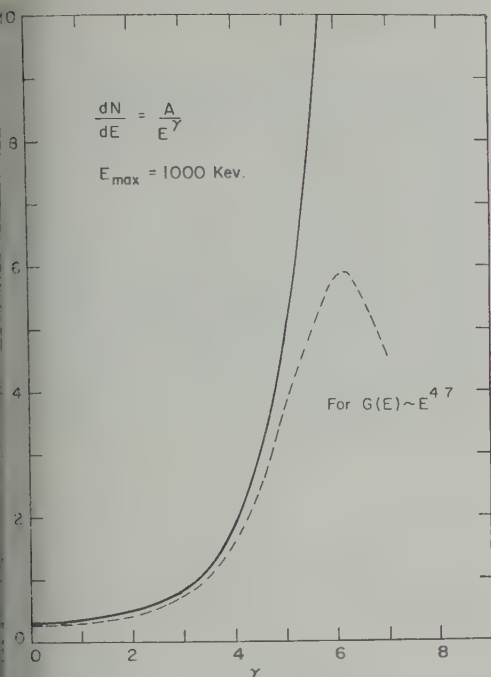


Fig. 25. The calculated relationship between the ratio Z and the exponent γ , assuming that the electron spectrum is $dN = AE^{-\gamma} dE$ ($E \leq 1$ Mev).

are not extremely sensitive to the shape of the low energy portion of the calibration curve except for high values of Z .

Thus the changes in spectrum at E_s (max), which occurred during recovery from geomagnetic storms in Figure 22, would be interpreted as a change from approximately E^{-3} to $\sim E^{-4}$ for the electron differential spectrum between August 20 and August 26, 1959.

If these assumptions were approximately correct, the dependence of Z on R_0 in Figure 20 would imply that the electron energy spectrum becomes 'flatter' with increasing range.

REFERENCES

Arnoldy, R., R. Hoffman, and J. R. Winckler, Measurements of the Van Allen radiation belts during geomagnetic storms, *Space Research*, First International Space Science Symposium, edited by H. K. Kallman Bijl, North Holland Publishing Company, pp. 877-896, Amsterdam, 1960a.
 Arnoldy, R. L., R. A. Hoffman, and J. R. Winckler, Observations of the Van Allen radiation regions during August and September 1959; 1, *J. Geophys. Research*, **65**, 1361-1375, 1960b.
 Bredin, J. B., L. F. Chase, W. L. Imhof, and D. J.

Knecht, Energy spectrum and angular distributions of electrons trapped in the geomagnetic field, *Tech. Rep. LMSD-895085*, 1961.
 Dessler, A. J., and R. Karplus, The gap in the electron component of the outer zone of the Van Allen radiation (abstract), *J. Geophys. Research*, **65**, 2486, 1960.
 Fan, C. Y., P. Meyer, and J. A. Simpson, Trapped and cosmic radiation measurements from Explorer VI, *Space Research*, First International Space Science Symposium, edited by H. L. Kallman Bijl, North Holland Publishing Company, pp. 951-966, Amsterdam, 1960.
 Fan, C. Y., P. Meyer, and J. A. Simpson, Equatorial pitch angle distributions of electrons in the outer radiation belt, Proceedings of the Second International Space Science Symposium, Florence, Italy, April, 1961.
 Farley, T. A., and A. Rosen, Charged particle variations in the outer Van Allen zone during a geomagnetic storm, *J. Geophys. Research*, **65**, 3494-3496, 1960.
 Freden, S. C., and R. S. White, Particle fluxes in the inner radiation belt, *J. Geophys. Research*, **65**, 1377-1383, 1960.
 Gold, T., Motions in the magnetosphere of the earth, *J. Geophys. Research*, **64**, 1219-1224, 1959.
 Herlofson, N., Diffusion of particles in the earth's radiation belts, *Phys. Rev. Letters*, **5**, 414, 1960.
 Hess, W. N., The radiation belt produced by neutrons leaking out of the atmosphere of the earth, *J. Geophys. Research*, **65**, 3107-3115, 1960.
 Hess, W. N., J. Killeen, C. Y. Fan, P. Meyer, and J. A. Simpson, The observed outer belt electron distribution and the neutron decay hypothesis, *J. Geophys. Research*, **66**, 2313-2314, 1961.
 Northrop, T. G., and E. Teller, Stability of the adiabatic motion of charged particles in the earth's field, *Phys. Rev.*, **117**, 215-225, 1960.
 O'Brien, B. J., J. A. Van Allen, F. E. Roach, and C. W. Gartin, Correlation of an auroral arc and a subvisible monochromatic 6300 Å arc with outer-zone radiation on November 28, 1959, *J. Geophys. Research*, **65**, 2759-2766, 1960.
 Parker, E. N., Effect of hydromagnetic waves in a dipole field on the longitudinal invariant, *J. Geophys. Research*, **66**, 693-708, 1961a.
 Parker, E. N., The Distribution of Trapped Particles in a Changing Magnetic Field, *J. Geophys. Research*, **66**, 2641-2652, 1961.
 Post, R. F., and W. A. Perkins, *Phys. Rev. Letters*, **5**, 85, 1961.
 Rosen, A., and T. A. Farley, Characteristics of the Van Allen radiation zones as measured by the scintillation counter on Explorer VI, *J. Geophys. Research*, **66**, 2013-2028, 1961.
 Rosen, A., T. A. Farley, and C. P. Sonett, Soft radiation measurements on Explorer VI earth satellite, *Space Research*, First International Space Science Symposium, edited by H. K. Kallman Bijl, North Holland Publishing Company, pp. 938-950, Amsterdam, 1960.
 Smith, E., and Rosen, A., The scientific results of the satellite: Explorer VI, in *Ballistic Missile and*

- Space Technology*, edited by D. P. LeGalley, Academy Press, p. 405, 1960.
- Sonett, C. P., E. J. Smith, D. L. Judge, and P. J. Coleman, Jr., Current systems in the vestigial geomagnetic field: Explorer VI, *Phys. Rev. Letters*, 4, 161-163, 1960.
- Van Allen, J. A., C. E. McIlwain, and G. H. Ludwig, Radiation observations with satellite 1958 epsilon, *J. Geophys. Research*, 64, 271-286, 1959a.
- Van Allen, J. A., C. E. McIlwain, and G. H. Ludwig, Satellite observations of electrons artificially injected into the geomagnetic field, *J. Geophys. Research*, 64, 877-891, 1959b.
- Vernov, S. N., A. Ye. Chudakov, P. V. Vakulov, and Yu. I. Logachev, Study of terrestrial corpuscular radiation and cosmic rays by the flight of a cosmic rocket, *Doklady Akad. Nauk SSSR*, 125, 5, 1959.
- Vernov, S. N., and A. Ye. Chudakov, Terrestrial corpuscular radiation and cosmic rays, *Space Research*, First International Space Science Symposium, edited by H. K. Kallman Bijl, North Holland Publishing Company, pp. 751-790, Amsterdam, 1960.
- Walt, M., L. F. Chase, Jr., J. B. Cladis, W. L. ... and D. J. Knecht, Energy spectra and altitude dependence of electrons trapped in the earth's magnetic field, *Space Research*, First International Space Science Symposium, edited by H. K. Kallman Bijl, North Holland Publishing Company, pp. 910-920, Amsterdam, 1960.

(Manuscript received June 29, 1961)

The Distribution of Trapped Particles in a Changing Magnetic Field

E. N. PARKER

*Enrico Fermi Institute for Nuclear Studies and Department of Physics
University of Chicago, Chicago 37, Illinois*

Abstract. The redistribution of charged particles in the mirror field $B(s, t) = B_0 T(t) \{1 + [s/a(t)]^{\nu(t)}\}$ is worked out for slow changes in $T(t)$, $a(t)$, and $\nu(t)$. It is found that increasing $T(t)$ gives a relatively greater particle density increase in the center of the field than deep in the mirrors $s \gg a(t)$. The mirror distance retracts like $1/T^{1/(\nu+2)}$. Decreasing $a(t)$ has the opposite effect. Field variations constrained to preserve $T(t) a^2(t)$ and $\nu(t)$ leave the form of the particle distribution unchanged, increasing the density everywhere by the same factor.

INTRODUCTION

Large changes are observed in the electron distribution and intensity in the outer Van Allen radiation belt during and after a geomagnetic storm. It is clear that irreversible processes must be present to account for some of the changes [Rosen and Farley, 1961; Fan, Meyer, and Simpson, 1961], but reversible changes associated with the slow storm deformations of the field must be present, too. To disentangle reversible and irreversible effects it is necessary to compute the reversible changes brought about by slow deformations of the field. Then, in principle, suitable measurements of the geomagnetic field in the outer Van Allen belt during a magnetic storm will make it possible to subtract out the reversible effects in a quantitative manner from the total storm changes in the electron intensity, and thereby determine precisely what irreversible changes take place.¹

When there is a slow change in the geomagnetic field, the individual particle motions will change in such a way as to satisfy the three constraints imposed by the three adiabatic invariants of the particle motion. Thus several effects may appear simultaneously, even for the simplest form of field variation. It is necessary to calculate the changes in the particle motion

in quantitative detail. For this purpose we choose the simple mirror field in which the field density $B(s, t)$ at a distance s along a line of force is given by

$$B(s, t) = B_0 T(t) \left\{ 1 + \left[\frac{s}{a(t)} \right]^{\nu(t)} \right\} \quad (1)$$

The line of force may be curved in any way we like, except that its radius of curvature must be large enough to satisfy the guiding center approximation. This field is appropriate to the analysis of particle density in any large-scale geomagnetic field change which does not produce local maxima in field density somewhere along the line of force. The magnitude of the field density is determined by $T(t)$ in terms of the characteristic field B_0 ; for convenience we let $T(0) = 1$. The scale of the field is given by $a(t)$, and the form by $\nu(t)$. All three of the functions $T(t)$, $a(t)$, and $\nu(t)$ may vary simultaneously, either because the field is changing slowly with time, or because the particle is drifting perpendicular to $B(s, t)$ into regions where T , a , and ν have different values, or both. Why $T(t)$, $a(t)$, and $\nu(t)$ vary with time is not essential to the discussion so long as we remember that the three parameters T , a , and ν are to be assigned the values appropriate to the location of the particle at the moment.² Thus we do not concern ourselves with the third invariant of the particle

¹ The original suggestion by Van Allen [1959], Van Allen and Frank [1959], and Van Allen, McIlwain, and Ludwig [1959], that the increase in radiation intensity is due to the capture of energetic solar particles has been ruled out, first by the arguments of Dessler and Karplus [1960] and second by the fact that sufficiently high radiation densities are not observed in interplanetary space.

² Presumably the changes in the field are due to the pressure of ionized gases against the field or trapped inside the field [Dessler and Parker, 1959]. The changes may be with respect to time at any given point in the field, or static changes around the region through which the particle drifts [Vestine, 1960].

motion, which has to do with how the particle distribution varies in the direction $\mathbf{B} \times \nabla B$. We consider in the next section only the first and second invariants, which tell us how the particles are deployed along the line of force on which they are instantaneously moving.

Omitting to write the argument of the functions $T(t)$, $a(t)$, and $\nu(t)$, it is readily shown that the volume enclosed by the flux tube with initial cross section A_0 at $s = 0$ is subsequently

$$V(t) = \frac{A_0 a}{T(t)} \frac{\pi/\nu}{\sin \pi/\nu} \quad (2)$$

It is evident that $V(t)$ is a monotonically decreasing function of $\nu(t)$. Later, we shall see how the particle density varies with $V(t)$.

INDIVIDUAL PARTICLE MOTION

The individual particle motion is determined uniquely, except for the phase of its cyclotron motion, by the transverse and longitudinal invariants while T , a , and ν are varying slowly with time. Consider a particle, trapped in the field (1), with pitch angle $\theta(s, t)$. The particle crosses $s = 0$ with pitch angle $\theta(0, t)$, oscillating between its two mirror points at $s = \pm s_m(t)$. Since $B(s, t)$ changes but little during one such oscillation, the transverse invariant yields

$$\frac{\sin^2 \theta(s, t)}{B(s, t)} \cong \frac{\sin^2 \theta(0, t)}{B(0, t)}$$

Thus if we write $S(s) \equiv 1 + (s/a)^\nu$, it follows that

$$\sin^2 \theta(s, t) = S(s) \sin^2 \theta(0, t) \quad (3)$$

and

$$S(s_m) \sin^2 \theta(0, t) = 1 \quad (4)$$

It also follows from the transverse invariant that the particle speed $w(t)$ is related to the initial particle speed w_0 by

$$w^2(t) \sin^2 \theta(0, t) = w_0^2 T(t) \sin^2 \theta_0 \quad (5)$$

where $\theta_0 \equiv \theta(0, 0)$. Finally, we have the condition from the longitudinal invariant that

$$\begin{aligned} w(t) \int_0^{s_m(t)} ds \cos \theta(s, t) \\ = w_0 \int_0^{s_m(0)} ds \cos \theta(s, 0) \end{aligned} \quad (6)$$

The conditions (5) and (6) are sufficient to

determine the subsequent particle motion described at $s = 0$ by $w(t)$ and $\theta(0, t)$, in terms of the initial motion described by w_0 and θ_0 . The motion along the line of force is described in terms of $w(t)$ and $\theta(0, t)$ and (4).

If we use (3) and (4) to express $\cos \theta$ in terms of $S(s)$, the integral in (6) is immediately evaluated to give

$$\begin{aligned} \frac{\Gamma(1/\nu + 1)}{\Gamma(1/\nu + 3/2)} \frac{w(t)a(t)}{S^{1/2}(s_m)} \left(\frac{s_m}{a} \right)^{1+\nu/2} \\ = \frac{\Gamma(1/\nu(0) + 1)}{\Gamma(1/\nu(0) + \frac{3}{2})} \frac{w_0 a_0}{S^{1/2}(s_m(0))} \left(\frac{s_m(0)}{a_0} \right) \end{aligned}$$

where $a_0 \equiv a(0)$. From (4) and (5) we have

$$\frac{w^2(t)}{S(s_m)} = T(t) \frac{w_0^2}{S(s_m(0))}$$

so that if

$$Q \equiv \frac{a_0}{a(t)} \frac{\Gamma(1/\nu(0) + 1) \Gamma(1/\nu + \frac{3}{2})}{\Gamma(1/\nu(0) + \frac{3}{2}) \Gamma(1/\nu + 1)}$$

then

$$\frac{s_m(t)}{a(t)} = \left\{ \left[\frac{s_m(0)}{a(0)} \right]^{\nu(0)+2} \frac{Q^2}{T(t)} \right\}^{1/(\nu+2)}$$

From (5) and (4) we have

$$\frac{w^2(t)}{w_0^2} = T(t) \frac{\sin^2 \theta_0}{\sin^2 \theta(0, t)} = T(t) \frac{S(s_m)}{S_0(s_m(0))}$$

where $S_0(s)$ is $1 + (s/a_0)^{\nu(0)}$. From (4) we have

$$\cot^2 \theta(0, t) = (s_m/a)^\nu$$

so that, with the aid of (9),

$$\begin{aligned} \cot^{2(\nu+2)/\nu} \theta(0, t) \\ = (Q^2/T) \cot^{2(\nu(0)+2)/\nu(0)} \theta_0 \end{aligned}$$

The pitch angle $\theta(s, t)$ is related to $\theta(0, t)$ (3), and

$$\sin^2 \theta(s, t) = S(s)/S(s_m)$$

so that

$$\begin{aligned} \frac{S(s_m)}{S_0(s_m(0))} \\ = \frac{S(s)/\sin^2 \theta(s, t)}{1 + \left(\frac{T}{Q^2} \right)^x \left[\frac{S(s)}{\sin^2 \theta(s, t)} - 1 \right]^{\nu(t)}} \end{aligned}$$

$$\frac{\nu(0)}{\nu(0) + 2} \quad y(t) = 1 + \frac{2(\nu(0) - \nu)}{\nu(\nu(0) + 2)}$$

is expression may now be put into (10) to give $w(t)$ as a function of the present pitch angle $\theta(s, t)$ and the initial velocity w_0 . Later on we will need the result that

$$\theta(s, 0) = S_0(s) \sin^2 \theta_0$$

$$1 + \left(\frac{T}{Q}\right)^x \left[\frac{S(s)}{\sin^2 \theta(s, t)} - 1 \right]^{\nu(t)} \quad (13)$$

The transit time $\tau(t)$ from $-s_m(t)$ to $+s_m(t)$ is

$$= \frac{2}{w(t)} \int_0^{s_m} \frac{ds}{\cos \theta(s, t)}$$

$$= \frac{2}{w(t)} \int_0^{s_m} \frac{ds}{[1 - S(s)/S(s_m)]^{1/2}}$$

$$\tau(t) = 2\pi^{1/2} a(t) \frac{\cot^{2/\nu-1} \theta(0, t)}{\sin \theta(0, t)}$$

$$\frac{\Gamma(1/\nu + 1)}{\Gamma(1/\nu + \frac{1}{2})} \quad (14)$$

GENERAL PARTICLE DISTRIBUTION

Variation of the particle distribution with s , and ν follows from the individual particle motions. Suppose that initially the number of particles per unit area crossing the $s = 0$ plane with pitch angles in $(\theta_0, \theta_0 + d\theta_0)$, during their oscillations between $\pm S_m(0)$, is $\Theta(0, \theta_0) d\theta_0$. Subsequently the number per unit area with pitch angle in $[\theta(0, t), \theta(0, t) + d\theta(0, t)]$, is $\psi(0, t) d\theta(0, t)$, where

$$\psi(0, t) = \Theta(0, \theta_0) T(t) \left[\frac{\partial \theta_0}{\partial \theta(0, t)} \right]_t \quad (15)$$

where the factor $T(t)$ takes account of the change of particle density due to the compression of the particles along with the field, where the subscript t indicates that the differentiation is to be carried out for a fixed value of t . Solving (11) for $\cot \theta_0$ it is readily seen that

$$\left[\frac{\partial \theta_0}{\partial \theta(0, t)} \right]_t = \frac{\nu(0)(\nu + 2)}{\nu(\nu(0) + 2)}$$

$$\frac{\sin^2 \theta_0 \cot^{\nu-1} \theta(0, t)}{\sin^2 \theta(0, t)} \cdot \left(\frac{T^{1/2}}{Q} \right)^x \quad (16)$$

If we let $\psi(s, t, \theta(s, t)) d\theta(s, t)$ represent the number of particles per unit volume at s with pitch angles in $[\theta(s, t), \theta(s, t) + d\theta(s, t)]$, then

$$\psi(s, t, \theta(s, t)) = \Theta(t, \theta(0, t))$$

$$\cdot \frac{S(s)}{w(t)\tau(t) \cos \theta(s, t)} \cdot \left[\frac{\partial \theta(0, t)}{\partial \theta(s, t)} \right]_{s,t} \quad (17)$$

The factor $S(s)$ represents concentration of the particles due to the converging lines of force of the magnetic field. The factor $w\tau \cos \theta$ arises from the fact that $ds/w\tau \cos \theta$ is the fraction of the total transit time τ spent in the interval $(s, s + ds)$. It is readily shown from (3) that

$$\left[\frac{\partial \theta(0, t)}{\partial \theta(s, t)} \right]_{s,t} = \frac{1}{S^{1/2}(s)} \frac{\cos \theta(s, t)}{\cos \theta(0, t)} \quad (18)$$

It follows immediately from (15) and (17) that

$$\psi(s, t, \theta(s, t)) = \Theta(0, \theta_0) \frac{S(s)T(t)}{w(t)\tau(t) \cos \theta(s, t)}$$

$$\cdot \left(\frac{\partial \theta_0}{\partial \theta(0, t)} \frac{\partial \theta(0, t)}{\partial \theta(s, t)} \right)_{s,t}$$

Applying (14), (16), and (18) yields

$$\psi(s, t, \theta(s, t))$$

$$= \frac{\Theta(0, \theta_0) S^{1/2}(s) T^g(t) \Gamma(1/\nu + \frac{3}{2})}{2\pi^{1/2} a(t) Q^{\nu(0)/(\nu(0)+2)} \Gamma(1/\nu + 1)}$$

$$\cdot \frac{\sin^2 \theta_0}{(1/\nu(0) + \frac{1}{2}) \sin^2 \theta(0, t) \cot^h \theta(0, t)}$$

where

$$g \equiv \frac{3\nu(0) + 4}{2(\nu(0) + 2)}, \quad h = \frac{2}{\nu} - (y - 1)$$

Hence

$$\psi(s, 0, \theta(s, 0))$$

$$= \frac{\Theta(0, \theta_0) S_0^{1/2}(s) \Gamma(1/\nu(0) + \frac{1}{2})}{2\pi^{1/2} a_0 \Gamma(1/\nu(0) + 1) \cot^{2/\nu(0)} \theta_0}$$

It follows then that

$$\psi(s, t, \theta(s, t)) = \psi(s, 0, \theta(s, 0)) T^{3/2}$$

$$\cdot \left[\frac{S(s)}{S_0(s)} \right]^{1/2} \cdot \frac{\sin^2 \theta_0}{\sin^2 \theta(0, t)} \quad (19)$$

after using (10) to eliminate $\cot \theta_0$ and (8) to eliminate Q . Finally, we use (10) and (12),

obtaining

$$\frac{\psi(s, t, \theta)}{S^{1/2}(s)} = \frac{\psi(s, 0, \theta(s, 0))}{S_0^{1/2}(s)} T^{3/2}(t) \cdot \frac{S(s)/\sin^2 \theta}{1 + F[S(s)/\sin^2 \theta - 1]^{\nu(t)}} \tag{20}$$

where we write θ in place of $\theta(s, t)$, and $F = (T/Q^2)^2$. The initial pitch angle $\theta(s, 0)$ is related to θ by (13). We have written (20) in the form given above to show that it is a solution of the Liouville equation

$$0 = \frac{\partial \psi}{\partial s} \cos \theta + \frac{\partial \psi}{\partial \theta} \frac{\sin \theta}{2S} \frac{dS}{ds} - \frac{\psi \cos \theta}{2S} \frac{dS}{ds} \tag{21}$$

for a quasi-stationary distribution ψ of particles along the field $S(s)$. The general solution of (21) may be written in the form

$$\frac{\psi(s, t, \theta)}{S^{1/2}} = G\left(\frac{\sin^2 \theta}{S(s)}\right) \tag{22}$$

where G is an arbitrary function. It is immediately seen from (21) that, if the initial distribution $\psi(s, 0, \theta(s, 0))/S_0^{1/2}$ is a function only of $\sin^2 \theta(s, 0)/S_0$, then it is a function only of $\sin^2 \theta/S(s)$. Hence, from (20), $\psi(s, t, \theta)/S^{1/2}(s)$ is a function only of $S(s)/\sin^2 \theta$.

In a mirror field, such as we are considering, we expect that the particle distribution will go to zero as $S(s) \rightarrow \infty$, suggesting that we will expand $G(x)$ in ascending powers of x ,

$$G(x) = \sum_{n=0}^{\infty} \frac{A_n}{x^{\alpha(n)}} \tag{23}$$

where $\alpha(n)$ is a monotonically increasing sequence of numbers, $\alpha(n + 1) > \alpha(n)$ for all n . When conditions are such that the initial particle distribution $\psi(s, 0, \theta(s, 0))/S_0^{1/2}(s)$ is expandable in powers of $\sin^2 \theta(s, 0)/S_0^{1/2}$, our treatment of the subsequent distribution as a function of $T(t)$, $a(t)$, and $\nu(t)$ is particularly simple. We carry out the expansion indicated by (23) and consider each term in the expansion separately. It is sufficient, therefore, to discuss (20) for the simple case that

$$\psi(s, 0, \theta(s, 0)) = \text{constant} \cdot \frac{\sin^\alpha \theta(s, 0)}{S_0^{(\alpha-1)/2}(s)} \tag{24}$$

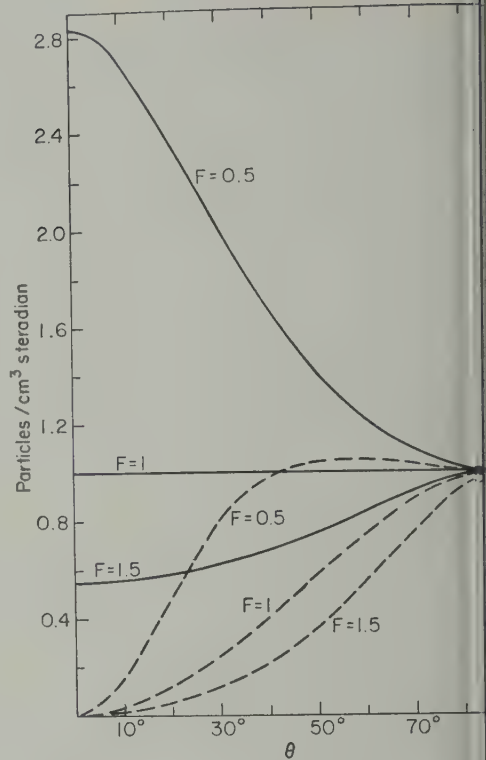


Fig. 1. Pitch angle distribution $\psi/\sin \theta$ at S plotted from (26) with $\gamma = 1$ for $F = 0.5, 1.0$ and normalized to unity at $\theta = \pi/2$. The curves are for $\alpha = 1$, and the broken curves $\alpha = 3$. When $S(s) \gg 1$ (and $\gamma = 1$) the form of pitch angle distribution is independent of F , responding then to the distribution for $F = 1$.

More complicated initial distributions are assumed to be expressible in terms of a superposition of these elementary distributions.

We normalize the distribution to give a particle density N at $s = 0$. Then

$$\begin{aligned} \psi(s, 0, \theta(s, 0)) &= \frac{N}{\pi^{1/2}} \frac{\Gamma(\alpha/2 + 1) \sin^\alpha \theta(s, 0)}{\Gamma(\alpha/2 + \frac{1}{2}) S_0^{(\alpha-1)/2}(s)} \end{aligned}$$

Subsequently the particle distribution is

$$\begin{aligned} \psi(s, t, \theta) &= \frac{N}{\pi^{1/2}} \frac{\Gamma(\alpha/2 + 1)}{\Gamma(\alpha/2 + \frac{1}{2})} \\ &\cdot \frac{Q^3 F^{3(\nu+2)/2\nu} S^{3/2}(s)/\sin^2 \theta}{\{1 + F[S/\sin^2 \theta - 1]^{\nu(t)}\}^{\alpha/2+1}} \end{aligned}$$

which is plotted in Figure 1 in comparison

0, θ). It is readily seen that, when $S \gg 1$,

$$t, \theta) \sim \frac{N}{\pi^{1/2}} \frac{\Gamma(\alpha/2 + 1)}{\Gamma(\alpha/2 + \frac{1}{2})} \frac{Q^3 F^{[6-\nu(\alpha-1)]/2\nu} \sin^{\alpha+(y-1)(\alpha+2)} \theta}{S^{(\alpha-1)/2+(y-1)(\alpha/2+1)}}$$

that the form of the θ dependence of ψ is changed if $y = 1$.

PARTICLE DENSITY

Consider how the particle density $N(s, t)$ is affected by changes in the field parameters $a(t)$, $\nu(t)$, and $\nu(t)$. If (25) is the initial particle distribution, the initial particle density is³

$$\begin{aligned} N(s, 0) &= \int_0^\pi ds \psi(s, 0, \theta) \\ &= N/S_0^{(\alpha-1)/2}(s) \end{aligned} \quad (27)$$

total number of particles in the flux tube of unit cross section at $s = 0$ is

$$\begin{aligned} N &= \int_0^\pi \frac{ds N(s, 0)}{S_0(s)} \\ N a_0 &= \frac{\Gamma(1/\nu(0) + 1) \Gamma[(\alpha + 1)/2 - 1/\nu(0)]}{\Gamma[(\alpha + 1)/2]} \end{aligned} \quad (28)$$

Consequently we know that the total number of particles is

$$\mathcal{N}(t) = T(t) \mathcal{N}(0) \quad (29)$$

As the particles move with the field. The density follows immediately from (26) as

$$\begin{aligned} N(s, t) &= \int_0^\pi d\theta \psi(s, t, \theta) \\ \frac{2N}{\pi^{1/2}} \frac{\Gamma(\alpha/2 + 1)}{\Gamma(\alpha/2 + \frac{1}{2})} T^{3/2} S^{3/2} & \\ \int_0^{\pi/2} \frac{d\theta}{\sin^2 \theta \{1 + F[S/\sin^2 \theta - 1]^{\nu(t)}\}^{\alpha/2+1}} & \end{aligned} \quad (30)$$

In order to study the effect upon $N(s, t)$ of variations of the field (1), we shall consider

the initial particle density as a function of field strength, $N(B)$, we need only expand N in descending powers of B . The coefficients in expansion of $N(B)$ become the coefficients of terms (25) in the expansion of ψ .

a number of special cases. With the substitution $z = \cot^2 \theta$ we have

$$\begin{aligned} N(s, t) &= \frac{N}{\pi^{1/2}} \frac{\Gamma(\alpha/2 + 1)}{\Gamma(\alpha/2 + \frac{1}{2})} T^{3/2} S^{3/2} \\ &\cdot \int_0^\infty \frac{dz}{z^{1/2} \{1 + FS^\nu[z + (1 - 1/S)^\nu]\}^{\alpha/2+1}} \end{aligned} \quad (31)$$

For the special case $\nu(t) = \nu(0)$ we have $y(t) = 1$ and

$$N(s, t) = N \frac{T^{3/2} S(s)}{F^{1/2} [1 - F + FS(s)]^{(\alpha+1)/2}} \quad (32)$$

Integrating $N(s, t)/S(s)$ from $-\infty$ to $+\infty$ yields (29). When $\nu(t) = \nu(0)$, we have $Q = a_0/a$ and $F = (Ta^2/a_0^2)^{\nu/(1+\nu)}$, so that we may also write

$$\begin{aligned} N(s, t) &= N \left(\frac{a_0}{a} \right)^3 \\ &\cdot \frac{F^{(\nu+3)/\nu} S(s)}{[1 - F + FS(s)]^{(\alpha+1)/2}} \end{aligned} \quad (33)$$

We see that if $T(t)$ and $a(t)$ vary in such a way that $Ta^2 = a_0^2$, then $F = 1$, and $N(s, t) = N(s, 0)(a_0/a)^3 = N(s, 0) T^{3/2}$. The density has the same dependence upon field strength $S(s)$ as initially, except for a uniform factor $T^{3/2}$, which in this case is proportional to the reciprocal of the volume given by (2). Thus compressing the field and shortening the scale at the same time affects the magnitude but not the form of the density dependence upon $S(s)$.

In general, a^2 and T will not vary in such a way as to maintain F constant. If we plot $N(s, t)/T^{3/2}(t)$ versus $S(s)$ from (32) for different values of F , we see how the form of the density distribution depends upon the scale $a(t)$ of the field. The results are shown in Figure 2 for $\alpha = 1, 3, 5$ and $\nu = 2$. Decreasing the scale of the field tends to spread the particles out along the field into the regions of stronger field. Plotting $N(s, t) a^3(t)/a_0^3$ versus $S(s)$ from (33) for various F shows how the form of the density distribution depends upon the field compression factor $T(t)$. The results are shown in Figure 3. We note that a uniform compression of the field yields a relative concentration of the particles toward low field. The bunching and spreading effects of $a(t)$ and $T(t)$ may be shown analytically

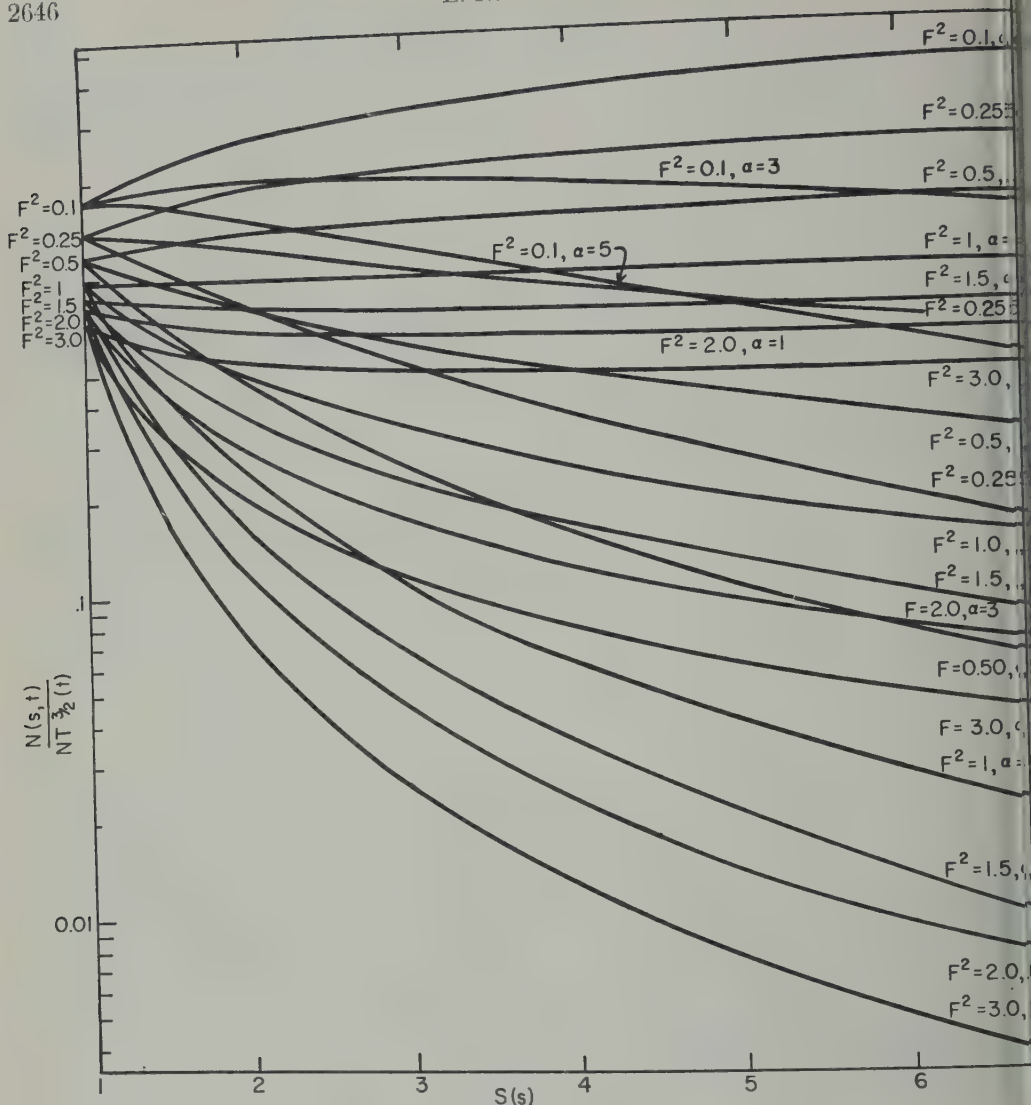


Fig. 2. Plot of $N(s, t)/N T^{3/2}(t)$ versus $S(s)$ for $\nu = 2$ from equation 32, showing the dependence of the particle density $N(s, t)$ upon the scale $a(t)$ of the field. The curves are given for various values of F^2 , to which $a(t)$ is related by $a(t) = a(0) F/T^{1/2}$.

from (32) and (33) by comparing the particle density at the center of the mirror $S = 1$ with the density deep in the mirror fields, $S \gg 1$. Then it is readily shown that

$$N(0, t) = N T^{(\nu+3)/(\nu+2)} (a_0/a)^{\alpha} \quad (34)$$

at the center of the mirror, and deep in the mirror field

$$N(s, t) \sim \frac{N}{S^{(\alpha-1)/2}} T^{\alpha} \left(\frac{a_0}{a}\right)^{\alpha(\alpha+2)} \quad (35)$$

where

$$r \equiv \frac{\nu(0) + 3}{\nu(0) + 2} \left[1 - \frac{(\alpha + 1)\nu(0)}{2(\nu(0) + 3)} \right]$$

The original density was $N/S^{(\alpha-1)/2}$ every-
where throughout the field. We see that at the
of the mirror where $S = 1$ the density inc-
reases with increasing $T(t)$ and with decreasing
At large S in the mirror field the densi-
creases even more rapidly with decreas-
and more slowly with increasing T . In fac-

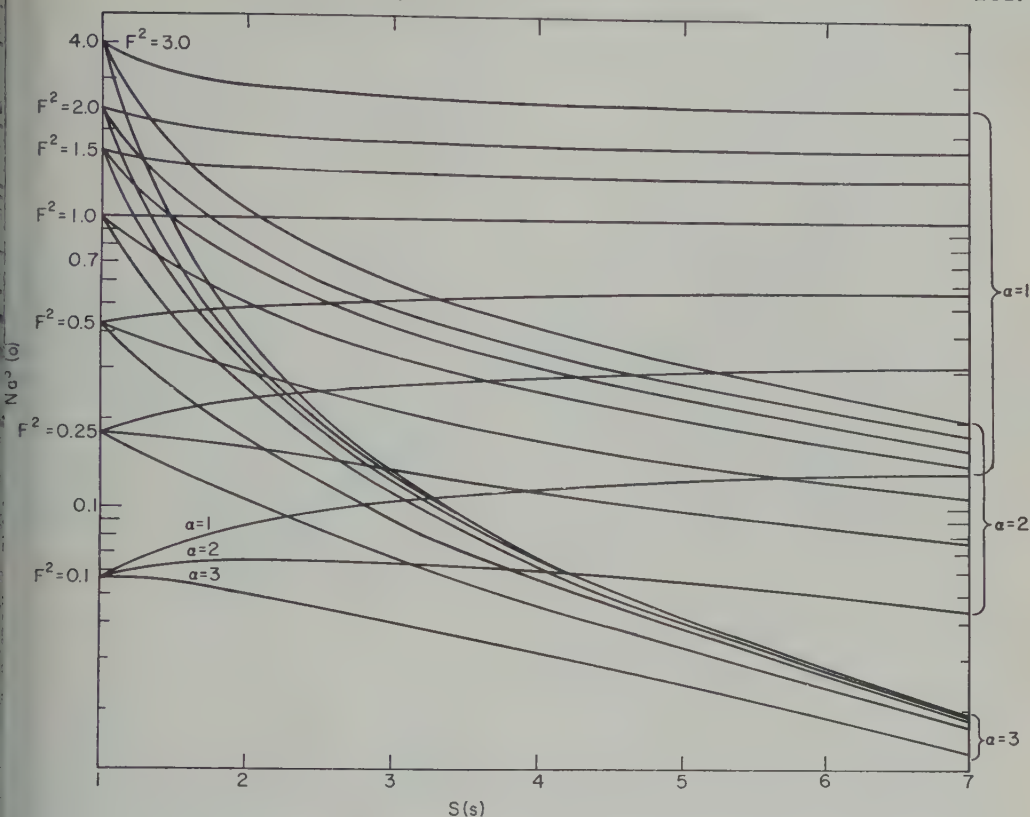


Fig. 3. Plot of $N(s, t) a^3(t) / Na^3(0)$ versus $S(s)$ for $\nu = 2$ from equation 33, showing the dependence of the particle density $N(s, t)$ upon the field magnitude $T(t)$. The curves are given for various values of F^2 , to which $T(t)$ is related by $T(t) = a_0^2 F^2 / a^2(t)$.

sufficiently large that $\nu(\alpha - 1) > 6$, the density at large S actually decreases with increasing T . Thus decreasing the scale $a(t)$ causes the particles to extend into regions of longer field, as the result of longitudinal acceleration; increasing the field magnitude $T(t)$ has the opposite effect, with proportionally more particles near the center of the mirror. Withdrawal of the particles from the region of strong field is given by (9). The mirror retracts like $1/T^{1/(\nu+2)}$.

Now suppose that we no longer impose the restriction that $y(t) = 1$. Then we have from that

at the center of the mirror. Where $S \gg 1$ we expand the integrand in ascending powers of $1/S$. Expanding first the quantity $(S/\sin^2 \theta - 1)^{\nu(t)}$ and then the entire denominator we obtain from (26) the particle distribution

$$\begin{aligned} \psi(s, t, \theta) = & \frac{N}{\pi^{1/2}} \frac{Q^{x(\alpha+2)}}{\Gamma(\alpha/2 + 1/2)} \\ & \cdot \frac{T^r \sin^{\alpha+2(y-1)} \theta}{S^{(\alpha-1)/2 + (y-1)(\alpha+2)/2}} \cdot \sum_{m=0}^{\infty} \frac{(-1)^m}{m!} \\ & \cdot \Gamma\left(\frac{\alpha}{2} + 1 + m\right) \cdot \left\{ \frac{1}{F} \left(\frac{\sin^2 \theta}{S} \right)^y \right. \\ & \left. + \sum_{n=1}^{\infty} \left(\frac{\sin^2 \theta}{S} \right)^n \frac{(-1)^n \Gamma(y+1)}{n! \Gamma(y+1-n)} \right\}^m \end{aligned}$$

$$N(0, t) = NT^{(\nu+3)/(\nu+2)} Q^{\nu/(\nu+2)} \frac{\Gamma\left[\frac{\alpha+1}{2} + \frac{\nu(0)-\nu}{\nu(0)(\nu+2)}\right] \Gamma\left[\frac{3}{2} - \frac{\nu(0)-\nu}{\nu(0)(\nu+2)}\right]}{\Gamma\left(\frac{\alpha+1}{2}\right) \Gamma\left(\frac{3}{2}\right)} \quad (36)$$

The density is the integral of ψ over θ . Taking only the leading term in the series, we obtain

$$N(s, t) \sim N \frac{T^r Q^{(\alpha+2)}}{S^{(\alpha-1)/2 + (y-1)(\alpha+2)/2}} \\ \cdot \frac{\Gamma(\alpha/2 + 1)}{\Gamma[(\alpha/2 + 1)y]} \\ \cdot \frac{\Gamma[(\alpha + 1)/2 + (y - 1)(\alpha + 2)/2]}{\Gamma[(\alpha + 1)/2]} \quad (37)$$

Comparing (36) and (37) with (34) and (35), we see that the density has the same dependence upon $T(t)$ and upon $a(t)$ (contained in Q) as when $\nu(t) = \nu(0)$. The density drops off with increasing S more rapidly when $y > 1$ ($\nu < \nu(0)$), so that at large S where (37) is valid the density is smaller by $S^{(y-1)(\alpha/2+1)}$. This is largely due to expansion of the field beyond $s = a$ when ν is decreased.

There is a change in the density apart from its dependence upon S due to $\nu \neq \nu(0)$, which arises from the gamma functions in (36) and (37). In (37) the change is of no particular importance, because it is dominated by $S^{(y-1)(\alpha/2+1)}$. It may not be unimportant in (36), and to illustrate it we suppose that $\nu(t)$ differs from $\nu(0)$ by only a very small amount, so that $\nu(t) = \nu(0)(1 + \xi)$, where $\xi \ll 1$. We note that

$$\Gamma(z + \epsilon)/\Gamma(z) \cong 1 + \epsilon\psi(z) + O(\epsilon^2)$$

where $\psi(z)$ is defined as $d \ln \Gamma(z)/dz$. The function $\psi(z)$ is monotonically increasing for $z > -1$, and is positive for $z > 0.4016$ [Jahnke and Emde, 1945]. It is readily shown that

$$Q \cong \frac{a(0)}{a(t)} \left\{ 1 - \frac{\xi}{\nu(0)} \left[\psi\left(\frac{1}{\nu(0)} + \frac{3}{2}\right) - \psi\left(\frac{1}{\nu(0)} + 1\right) \right] \right\} + O(\xi^2)$$

so that Q is a decreasing function of ξ . We find ultimately that

$$N(0, t) = NT^{(y+3)/(y+2)} \left(\frac{a_0}{a}\right)^{\nu/(y+2)} \\ \cdot \left\{ 1 - \frac{\xi}{\nu(0) + 2} \left[\psi\left(\frac{1}{\nu(0)} + \frac{3}{2}\right) - \psi\left(\frac{1}{\nu(0)} + 1\right) + \psi\left(\frac{\alpha}{2} + \frac{1}{2}\right) - \psi\left(\frac{3}{2}\right) \right] \right\} \quad (38)$$

Whether $N(0, t)$ is an increasing or decreasing function of ξ depends upon α and $\nu(0)$. If $\nu(0) = 2$, as is approximately the case for geomagnetic field, then $N(0, t)$ is an increasing or decreasing function according as α is less or greater than 1.190.

OBSERVED QUANTITIES

Having illustrated the manner of variation of the particle density with $T(t)$, $a(t)$, and S , we turn now to such quantities as particle energy density and omnidirectional intensity which are probably more closely related to what is observed by counter experiments in space. Particle energy density is the basis for the calculation of magnetic effects of trapped particles [Parker, 1957; Dessler and Parker, 1959]; particle counting rate measure omnidirectional intensity with energy-dependent efficiency.

Consider the quantity

$$J(s, t, w_0, \alpha, n) \equiv \int_0^\pi d\theta \psi(s, t, \theta) w^n(t)$$

where $\psi(s, t, \theta)$ is (26), so that it represents the subsequent distribution of particles in phase space with speed w_0 and density N at the center of the mirror with distribution (25). The function $J(s, t, w_0, \alpha, 1)$ represents the omnidirectional intensity, and $CJ(s, t, w_0, \alpha, n)$ represents the counting rate of an omnidirectional detector which measures the omnidirectional intensity with efficiency Cw^{n-1} . The energy density is $\frac{1}{2} MJ(s, t, w_0, \alpha, 2)$ for particles of mass m . If more than one type of particle is present, and if more than one initial velocity is present, the total omnidirectional intensity, energy density, etc., are obtained by summing over $J(s, t, w_0, \alpha, n)$ for the different kinds of particles and integrating over the velocity spectra.

From (26), and from (10) and (12), we obtain

$$J(s, t, w_0, \alpha, n) = \frac{2Nw_0^n}{\pi^{1/2}} \\ \cdot \frac{\Gamma(\alpha/2 + 1)}{\Gamma(\alpha/2 + 1/2)} T^{(n+3)/2} S^{(n+3)/2} \\ \cdot \int_0^{\pi/2} \frac{d\theta}{\sin^{n+2}\theta [1 + F(S/\sin^2\theta - 1)^y]^{(\alpha+1)/2}}$$

It is immediately obvious, as it was earlier for $n = 0$, that the form of J as a function

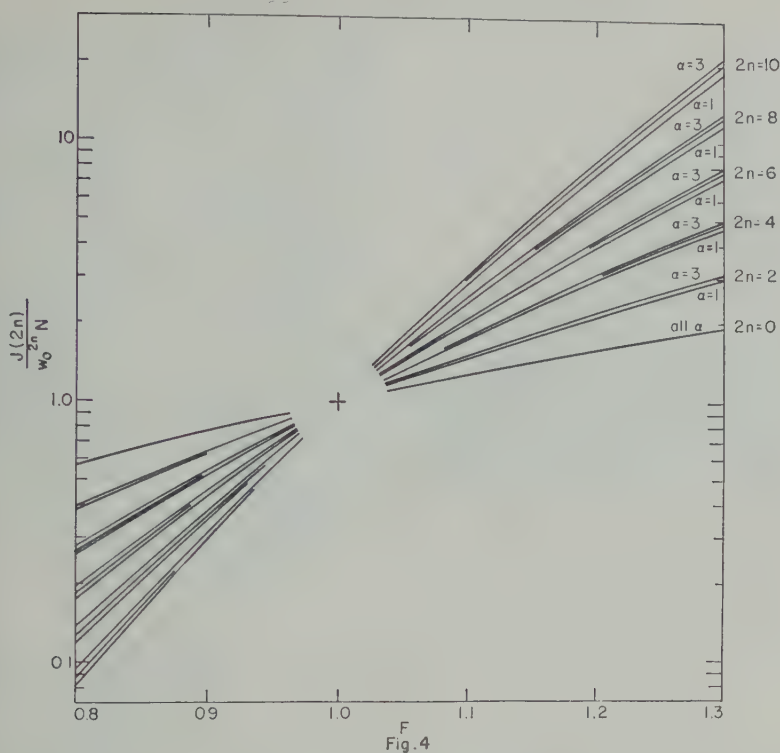


Fig. 4. Plot of $J(s, t, w_0, \alpha, 2n)$ at $S = 1$ for $\nu = 2$ and $n = 0, 1, \dots, 5$ and $\alpha = 1, 2, 3$ computed from equation 43. Linear interpolation to nonintegral values of n is permissible. $T(t) = a_0^2 F^2/a^2(t)$.

is invariant to changes in the field which F and y constant, i.e., which satisfy $a^2(t) = a_0^2$ and $\nu(t) = \nu(0)$. For such

$J(s, t, w_0, \alpha, n) = T^{(n+3)/2} J(s, 0, w_0, \alpha, n)$ which is proportional to $1/V^{1+n/3}$, where V is given by (2).

For the special case that $y = 1$ the integral can be evaluated in closed form if both α and n are integers and α is odd or n is even. If α is an integer, then we let $\mu = (1 - F)^{1/2} \cos \theta$. The numerator of the integrand becomes a polynomial in μ of order $(\alpha - 1)/2$. Then, for instance, the omnidirectional intensity $J(s, w_0, \alpha, 1)$ becomes

$J(s, w_0, 1, 1)$

$$= \frac{N w_0}{\pi^{1/2}} \frac{\Gamma(\alpha/2 + 1)}{\Gamma(\alpha/2 + \frac{1}{2})} \frac{T^{2-x} Q^{2x} S}{1 - F + FS} + \frac{FS}{2(1 - F)^{1/2}(1 - F + FS)^{1/2}}$$

$$\cdot \ln \frac{[(1 - F + FS)^{1/2} + (1 - F)^{1/2}]^2}{FS} \quad (41)$$

for $\alpha = 1$, and

$$J(s, t, w_0, 3, 1)$$

$$= \frac{N w_0}{2\pi^{1/2}} \frac{\Gamma(\alpha/2 + 1) T^{2-x} Q^{2x} S}{\Gamma(\alpha/2 + \frac{1}{2})(1 - F)(1 - F + FS)^2} \cdot \left\{ (1 - F - \frac{1}{2}FS) + \frac{FS(1 - F + \frac{1}{4}FS)}{(1 - F)^{1/2}(1 - F + FS)^{1/2}} \cdot \ln \frac{[(1 - F + FS)^{1/2} + (1 - F)^{1/2}]^2}{FS} \right\} \quad (42)$$

for $\alpha = 3$, etc., where again $F = (T/Q^2)^x$. If n is an even integer we let $u = \chi^{-1} \cot^2 \theta$, where $\chi = (1 - F + FS)/FS$. Then the integrand becomes a polynomial in u of order $n/2$.

We obtain

$$J(s, t, w_0, \alpha, 2n) = \frac{N w_0^{2n}}{\pi^{1/2}} \frac{F^{(2n+3-x)/2x} S^{n+1}}{(1 - F + FS)^{(\alpha+1)/2+n}} \left(\frac{a_0}{a}\right)^{2n+3} \\ \cdot \frac{\Gamma(\alpha/2 + 1)}{\Gamma(\alpha/2 + 1 + n)} \sum_{m=0}^{\infty} \frac{\chi^m}{m!} \\ \cdot \Gamma(m + \frac{1}{2}) \frac{\Gamma(\alpha/2 + \frac{1}{2} + n - m)}{\Gamma(\alpha/2 + \frac{1}{2})} \quad (43)$$

Deep in the mirror field it is possible to obtain an asymptotic expression for J for general α , n , and y . Expanding the integrand of (40) in ascending powers of $\sin^2 \theta/S$, and keeping only the leading term, just as we did in obtaining (37), we obtain

$$J(s, t, w_0, \alpha, n) \sim N w_0^n \frac{T^{r+n/(\nu(0)+2)} Q^{x(\alpha+2+n)}}{S^{(\alpha-1)/2+(y-1)(\alpha+2+n)/2}} \\ \cdot \frac{\Gamma(\alpha/2 + 1) \Gamma(\alpha/2 + \frac{1}{2} + (y-1)(\alpha+n+3)/2)}{\Gamma(\alpha/2 + 1 + (y-1)(\alpha+2+n)/2) \Gamma(\alpha/2 + 1/2)}$$

The dependence upon S deviates from the initial $S^{(\alpha-1)/2}$ when $\nu(t) \neq \nu(0)$ more sharply the larger n is. If $\nu(t) = \nu(0)$, then

$$J(s, t, w_0, \alpha, n) \sim \frac{N w_0^n T^{r+n/(\nu+2)}}{S^{(\alpha-1)/2}} \left(\frac{a_0}{a}\right)^{x(\alpha+2+n)} \quad (45)$$

which is more sensitive to $T(t)$ and to $a(t)$ the larger n is, because J then responds more to particle acceleration.

APPLICATION TO TRAPPED PARTICLE OBSERVATION

Consider what aspects of the observed trapped particle density are relevant to a theoretical analysis of the distribution in the geomagnetic field. The individual particle oscillates along a line of force between mirror points and drifts slowly around the earth, generating a shell which is open at each end beyond the mirror points [Northrup and Teller, 1959].

The particle density along any line of force in the given shell is intimately related to the pitch angle distribution of the particles, so that, given either, the other may be deduced directly through (22). The theory given above, of the

particle density distribution, demonstrates that the initial density distribution along the line of force will change with the field parameters $T(t)$, $a(t)$, and $\nu(t)$. In the simplest cases might choose to ignore changes in $\nu(t)$, put $\nu(t) = \nu(0) \cong 2$ to simulate the geomagnetic field. Then the changes in the particle distribution along a field line may be resolved into components. The first component consists of field changes which preserve the product $a^2(t)$, for which J changes everywhere by a factor $T^{(n+3)/2}$ (see equation 40) and the mirror positions remain at the same field density. The second component consists of changes in $a(t)$ alone, leaving $a(t)$ fixed. Increasing $T(t)$ to increase the particle intensity at the equator ($S = 1$) more than at high latitudes ($S > 1$).

with a general rise of mirror points in the field. It is obvious that any given changes in $T(t)$ and $a(t)$ can be represented by a combination of these two deformations.

To facilitate the analysis of observations we point out that the two essential features of change in density distribution along a line of force are the factors by which the density changes at the equator ($S = 1$) and at high latitude ($S \gg 1$). Suitable values of $T(t)$ and $a(t)$ can fit any two such factors. In Figure 4 we plot J from (43) at $S = 1$ as a function of F for the special case that $\nu = 2$ ($F = 1$). In Figure 5 we give the factor, from equation 45, for $\nu = 2$, by which J changes at latitude S as a function of F .

In applying these theoretical results to actual observations it must be remembered first that they are nonrelativistic, and hence are reliable above about 0.5 Mev (for electrons). (In a recent paper by Coleman [1961] the relativistic case is worked out for a uniform field.) Second, the reader is cautioned that the theoretical results have been worked out for particles of a single velocity w , whereas the actual Van Allen electrons exhibit a wide range of velocities. Therefore, comparing J with the counting

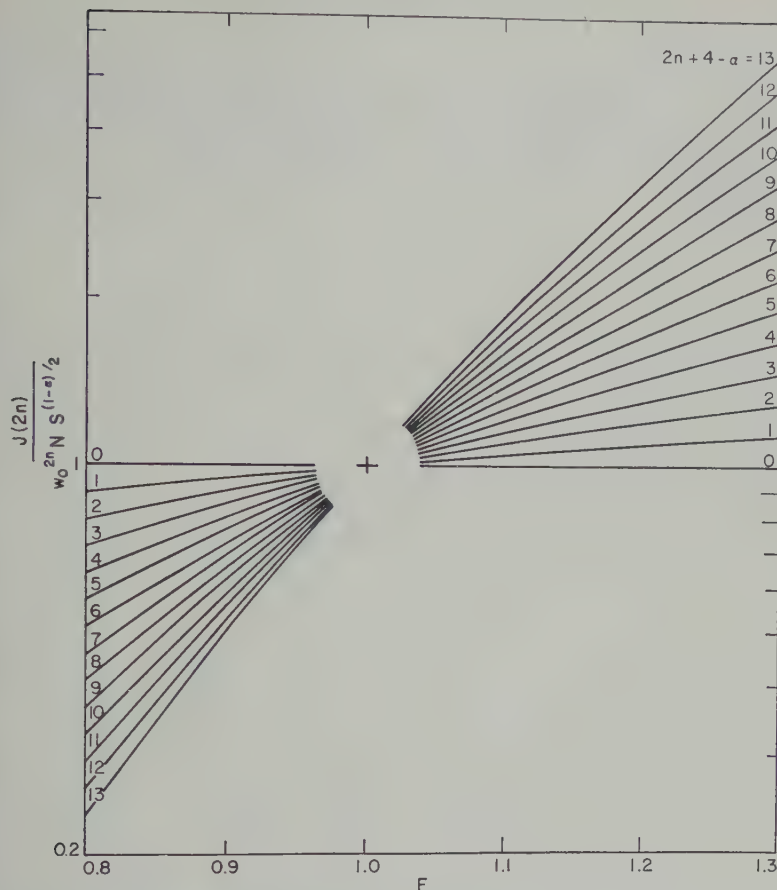


Fig. 5. Plot of $J(s, t, w_0, \alpha, 2n)$ for $S \gg 1$ when $a(t) = a_0$, $\nu = 2$, and $n = 0, 1, \dots, 5$, and $\alpha = 1, 2, 3$ computed from equation 45. Linear interpolation to nonintegral α and n is permissible. $T(t) = F^2$.

an actual detector which does not resolve particle energies involves the implicit assumption that the pitch angle distribution is the same at all energies.

Acknowledgment. This research is supported by the Office of Scientific Research, Air Research and Development Command, contract AF 18 (60)-666, and the National Aeronautics and Space Administration under grant NASA-NsG-96-60.

REFERENCES

Leeman, P. J., The effects of betatron accelerations upon the intensity and energy spectrum of magnetically trapped particles, *J. Geophys. Research*, **66**, 1351, 1961.
 Essler, A. J., R. Karplus, Some properties of the Van Allen radiation, *Phys. Rev. Letters*, **4**, 71, 1960.
 Essler, A. J., and E. N. Parker, Hydromagnetic

theory of geomagnetic storms, *J. Geophys. Research*, **64**, 2239-2252, 1959.
 Fan, C. Y., P. Meyer, and J. A. Simpson, Structure and dynamics of the outer radiation belts, *J. Geophys. Research*, **66**, 2607-2640, 1961.
 Jahnke, E., and E. Emde, *Tables of Functions*, Dover Publications, New York, 1945.
 Northrop, T. G., and E. Teller, Stability of the adiabatic motion of charged particles in the earth's field, *Univ. of Calif. Lawrence Radiation Lab., UCRL 5615*, 1959.
 Parker, E. N., Newtonian development of the dynamical properties of ionized gases of low density, *Phys. Rev.*, **107**, 924, 1957.
 Parker, E. N., Geomagnetic fluctuations and the form of the outer zone of the Van Allen radiation zone, *J. Geophys. Research*, **65**, 3117-3130, 1960.
 Parker, E. N., Effect of hydromagnetic waves in a dipole field on the longitudinal invariant, *J. Geophys. Research*, **66**, 693-708, 1961.
 Rosen, A., and T. Farley, 1961, Characteristics of the Van Allen radiation zones as measured by

- the scintillation counter on Explorer VI, *J. Geophys. Research*, 1961.
- Van Allen, J. A., The geomagnetically trapped corpuscular radiation, *J. Geophys. Research*, 64, 1683-1689, 1959.
- Van Allen, J. A., and L. A. Frank, Radiation measurements to 658,300 km with Pioneer IV, *Nature*, 184, 219, 1959.
- Van Allen, J. A., C. E. McIlwain, and G. H. Ludwig, Radiation observations with satellite 1959, *J. Geophys. Research*, 64, 271-286, 1959.
- Vestine, E. H., Polar auroral, geomagnetic, and ionospheric disturbances, *J. Geophys. Research*, 65, 360-362, 1960.

(Manuscript received June 29, 1961.)

Forbush Decrease of the Flux of Heavy Primary Nuclei of Cosmic Rays on May 12 and July 12, 1959

S. BISWAS¹

*School of Physics, University of Minnesota
Minneapolis 14, Minnesota*

Abstract. The time variation of the flux of heavy nuclei with $Z \geq 3$ of primary cosmic radiation has been measured during seven balloon flights on May 6-12, 1959, and on July 10-12, 1959. Six of these flights were made from Minneapolis and one from Churchill, Canada. During a quiet day on May 6, 1959, the flux of these nuclei at the top of the atmosphere was 21.2 ± 2.3 particles/m²-sec-sr. After the Forbush decrease on May 12, 1959, the value was 10.0 ± 1.4 particles/m²-sec-sr, showing a reduction of 53 ± 9 per cent of the primary heavy nuclei flux. On July 10-11, the flux of $Z \geq 3$ nuclei at the top of the atmosphere was measured as 17.5 ± 1.7 particles/m²-sec-sr before the Forbush decrease and 10.0 ± 1.8 particles/m²-sec-sr after the Forbush decrease on July 12. The reduction was 43 ± 9 per cent during this event. The Deep River neutron monitor showed a reduction of counting rates of 12 and 7.5 per cent respectively during these two Forbush decrease events.

Introduction. A 3+ solar flare occurred on May 10, 1959, from 2055 to 2600 UT. A sudden commencement of a geomagnetic storm occurred at 2320 UT on May 11, and there was a large Forbush decrease of intensity of neutron monitors at sea level. During this period after the magnetic storm, high intensity solar protons was observed at balloon altitudes at Minneapolis [Ney, Winckler, and Reier, 1959; Winckler and Bhavsar, 1960]. The intensity of solar α -particles was also measured at Minneapolis during this period (Reier, to be published). The measurements of flux of heavy nuclei during this period are reported in this paper. It is found that at atmospheric depth of about 10 g/cm² the flux of heavy nuclei was significantly lower than its normal value. This shows that the heavy nuclei of solar flare origin were absent in the flux of heavy nuclei detected at balloon altitude, and the measured flux of these nuclei represented the intensity of galactic cosmic radiation which had suffered Forbush decrease. The magnitude of the Forbush decrease of heavy nuclei of primary cosmic radiation during this event has been obtained from the measurements of the variation of the intensity near the top of

the atmosphere in a number of flights before and after the magnetic storm.

Similar measurements were made during the solar flare event of July 10, 1959. A class 3+ solar flare occurred from 0210 to >0900 UT on July 10. The sudden commencement of the magnetic storm occurred at 1623 UT on July 11. This was accompanied by a large Forbush decrease of intensity of neutron monitors at sea level [Winckler, Bhavsar, and Peterson, 1961].

The time variation of the intensity of heavy nuclei of primary cosmic rays during these two events has been studied by measuring the flux of heavy nuclei in seven emulsion stacks that were flown in seven balloon flights. Six of these balloon flights were made from Minneapolis and one from Churchill, Canada. The relevant data for these balloon flights are given in columns 1-4 in Table 1. The counting rates of the sea-level neutron monitor at Deep River, Canada, during this period are shown in the lower part of Figure 1 [Carmichael and Steljes, 1959].

In this study of time variation, the total flux of nuclei with $Z \geq 3$ has been used, with the advantage that these nuclei can be easily resolved from α -particles without ambiguity. Each of these nuclei detected at balloon altitude is of primary origin or a daughter of a heavier primary particle. Further, the experimentally observed absorption mean free path in air of nuclei with $Z \geq 3$ is as long as 45 g/cm² of air;

¹On deputation from Tata Institute of Fundamental Research, Bombay.

TABLE 1. Summary of Balloon Flights and the Measurements of the Flux of Heavy Nuclei

Flight* No.	Date 1959	Time at Ceiling UT	Balloon Ceiling, g/cm ²	Charge Group	Energy, Mev/ nucleon	Flux at 0 g/cm ² of air, particles/m ² ·sec·sr		Reduced of Flux
						Before Mag- netic Storm	After Mag- netic Storm	
IGC-7	5/6	0540-2106	14.2	$Z \geq 3$	>300	21.2 ± 2.3		
BP473	5/12	0920-1337	13.5	$Z \geq 3$	>310		11.1 ± 2.0	
BP475	5/12	1340-1752	12.0	$Z \geq 3$	>280		8.6 ± 1.9	
BP473 } BP475 }	5/12	0920-1752	12.8	$Z \geq 3$	>300		10.0 ± 1.4	53 ± 9
IGC-12	7/10	0930-2130	10.7	$Z \geq 3$	>260	17.0 ± 2.1		
IGC-13	7/11	0500-1925	12.0	$Z \geq 3$	≥ 275	18.2 ± 2.9		
IGC-12 } IGC-13 }	7/10	0930-						
	7/11	1925	11.4	$Z \geq 3$	>270	17.5 ± 1.7		
M-2	7/12	0720-1835	6.4	$Z \geq 3$	>200		10.0 ± 1.8	43 ± 9
IGC-12 } IGC-13 }	7/10	0930-						
	7/11	1925	11.4	$Z \geq 10$	>420	2.5 ± 0.5		
479N	7/12	1300-2345	2.8	$Z \geq 10$	>420		0.83 ± 0.12	67 ± 1

* All flights were launched from Minneapolis, except Flight 479N which was launched from Churchill, Canada.

as a result, the total flux of heavy nuclei of $Z \geq 3$ decreases rather slowly with atmospheric depth.

Experimental procedure. Each of the emulsion stacks in this investigation was composed of 24 Ilford G5 emulsions of size 10 by 10 by 0.06 cm. The emulsions were scanned along a line 6-cm long and 5 mm from the top edge of the emulsions. The scan was made for all particles crossing this line having ionization >8 times minimum ionization, zenith angle $<60^\circ$, and projected length >4 mm per plate. Particles having zenith angle $\leq 45^\circ$ and projected length >5 mm per plate were selected for measurements and were followed through the stack until they ended, interacted, or left the emulsion stack. The 'bowing' of emulsions in the region 0-1 cm from the top edge was measured and found to be negligible for all stacks except one. For this stack (IGC-7) proper geometrical criteria for selection of tracks were used so as to overcome the effect of bowing. Scanning efficiency was checked in some of the stacks and found to be close to 100 per cent for heavy nuclei of $Z \geq 3$. The unprocessed thickness of each emulsion sheet of this stack was obtained from the measurements of weights and thicknesses of individual emulsion sheets before processing.

The particles were identified as protons,

α -particles, and heavy nuclei of $Z \geq 3$ by ionization-range, δ -ray density-range, or δ -ray density-scattering measurement. Scattering measurements were made with cell sizes of 100 and 500 μ , and particles with $Z = 3$ or 4 were unambiguously separated from α -particles. Calibration of δ -ray density with Z was done for $Z = 2-8$ in the usual way from charts indicating breakups of relativistic heavy nuclei. In the three stacks IGC-12, 13, and 479N, the flux of nuclei of $Z \geq 10$ was measured. In the three stacks δ -ray density calibration was done for $Z = 8-12$.

The emulsion stack 479N, flown from Churchill, Canada, had very high density of solar proton tracks near the top edge, and line scanning near the top edge was not possible. Therefore the two side plates were 'area-scanned' for particles entering the stack having ionization >8 times minimum ionization. By ionization-range method, the energy spectrum of $Z \geq 10$ was measured in the energy interval 200-500 Mev/nucleon (to be published). From this the integral flux of $Z \geq 10$ of energy >420 Mev/nucleon was obtained.

In order to extrapolate the measured flux to flight altitude to the top of the atmosphere, knowledge of the absorption mean free path of $Z \geq 3$ nuclei is necessary. This is obtained in the following way: The flux of heavy primary

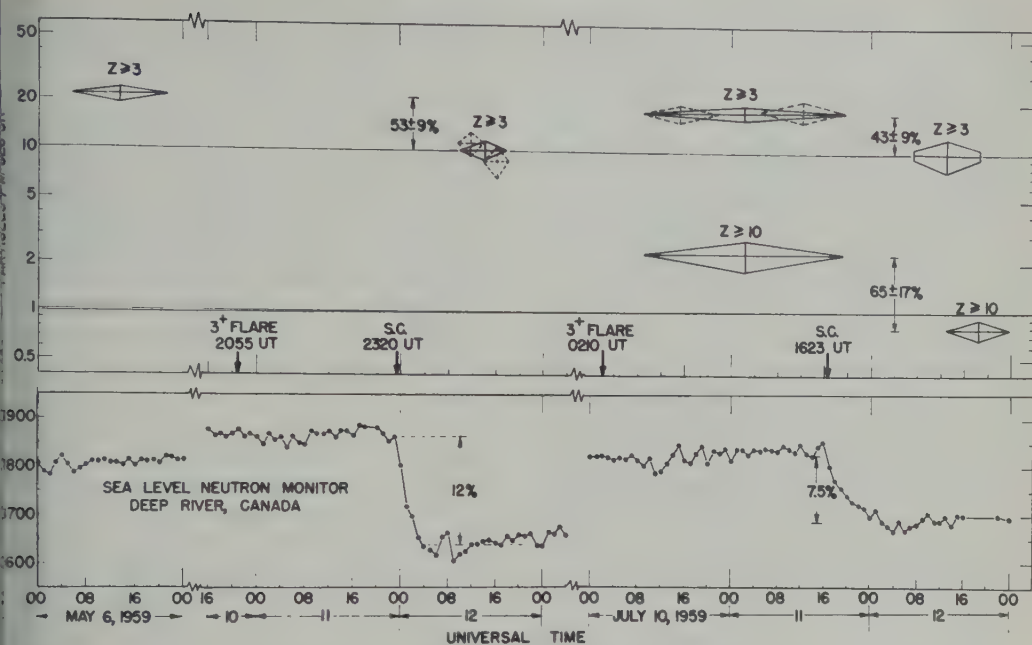


Fig. 1. The upper figure shows the time variation of the vertical flux of heavy nuclei of $Z \geq 3$ (and $Z \geq 10$) extrapolated to the top of the atmosphere as measured in seven balloon flights during May 6-12, 1959, and July 10-12, 1959. The flux of heavy nuclei with $Z \geq 3$ suffered Forbush decrease of 53 ± 9 per cent on May 12 and 43 ± 9 per cent on July 12, and the heavy nuclei of $Z \geq 10$ a decrease of 65 ± 17 per cent on July 12. The lower figure shows the counting rates of the Deep River neutron monitor during the same periods.

lei of energy >1.45 bev/nucleon during et days was assumed to be almost constant r sunspot minimum during 1954 to early pruary 1956. This is indicated by measure- ts of α -particle flux of energy >1.45 bev/ leon. The values of the measured flux of ≥ 3 nuclei of energy >1.45 bev/nucleon for ith angle $<60^\circ$ at flight altitude during this iod, as obtained by Waddington [1957], ter, Debenedetti, Garelli, Quassiat, Tallone, l Vigone [1958], and Appa Rao, Biswas, iel, Neelakantan, and Peters [1958], have n plotted against the corresponding atmos- ric depths (ranging from 10 to 35 g/cm² of). The absorption mean free path for $Z \geq 3$ lei is thus obtained as 45 ± 11 g/cm² of air. may be noted that while there are disagree- ts in these investigations about the ratio LiBeB nuclei to $Z \geq 6$ nuclei extrapolated the top of the atmosphere, there is good cement in the measured flux values of $Z \geq 3$ lei. The absorption mean free path in air

for $Z \geq 3$ nuclei turns out to be about the same as for α -particles. This arises from the well-known fact that interaction of heavy nuclei leads to a fragment of lower charge with high probability.

In this investigation the measured vertical flux values of $Z \geq 3$ nuclei have been extrapolated to the top of the atmosphere using absorption mean free path of 45 g/cm² of air. It may be remarked that since most of the measurements are made close to 10 g/cm² of air, the relative values of the flux of $Z \geq 3$ are almost independent of the absorption mean free path used for extrapolation. For nuclei with $Z \geq 10$, an absorption mean free path of 30 g/cm² of air has been used. The flux values have been corrected for particles entering the stack during the ascent of the balloon.

Results and discussion. The flux values of heavy nuclei extrapolated to the top of the atmosphere are given in Table 1 and in Figure 1. Flight IGC-7 was made during the period of no

solar or magnetic activity, and this gives the normal flux of 21.2 ± 2.3 particles/m²sec-sr of galactic cosmic radiation at this period of the solar cycle. The lower limit of energy of heavy nuclei is determined by the energy of the particle capable of penetrating the residual air above the balloon and the amount of emulsion above the scan line. This minimum energy is a function of the charge of the incident nuclei. For $Z \geq 3$ nuclei the average charge is about $Z = 7$, and the minimum energy for $Z_{av} = 7$ is shown in column 6 of Table 1. The flux values in the two flights BP473 and BP475 are the same within experimental errors, and therefore these have been combined to obtain the value of 10.0 ± 1.14 particles/m²sec-sr during the magnetic storm. This shows that during the Forbush decrease on May 12, 1959, the flux of nuclei with $Z \geq 3$ suffered a reduction of 53 ± 9 per cent. The Deep River neutron monitor showed a decrease of 12 per cent during this event.

During the July 10 event, before the magnetic storm, the flux values in flights IGC-12 and IGC-13 are the same, and therefore the average value of the prestorm flux has been obtained as 17.5 ± 1.7 particles/m²sec-sr. This value is about the same as that measured in flight IGC-7, indicating the absence of heavy nuclei of energy >280 Mev/nucleon from the solar flare of July 10, 1959. After the magnetic storm the flux of these nuclei was reduced to 10.0 ± 1.8 particles/m²sec-sr. This shows that there was a reduction of primary flux of $Z \geq 3$ nuclei of 43 ± 9 per cent during the Forbush decrease on July 12, 1959. The Deep River neutron monitor showed a Forbush decrease of 7.5 per cent during this period.

The flux of $Z \geq 10$ nuclei was also measured before the magnetic storm in flights IGC-12 and 13, and after the magnetic storm on flight 479N. This shows that the flux of $Z \geq 10$ nuclei of energy > 420 Mev/nucleon was reduced by 65 ± 17 per cent during the Forbush decrease. This value, with large statistical error, is consistent with the value of 43 ± 9 per cent obtained for $Z \geq 3$ nuclei.

Similar results of large reduction of flux of primary α -particles and heavy nuclei during Forbush decrease events have been published [McDonald, 1957; Biswas, Lavakare, Neelakantan, and Shukla, 1960; Meyer, 1960]. During the Forbush decrease of May 1959, McDonald

and Webber [1960] measured the different rigidity spectrum of galactic protons and α -particles at the low rigidity end on May 16 June 2. On May 16 the Deep River neutron monitor showed partial recovery from the Forbush decrease of May 12, but it was 8 per cent below the normal value reached June 2. Their results show that protons and α -particles of the same rigidity participated in the same way during the Forbush decrease. The flux of both protons and α -particles of rigidity $R > 1.4$ bev on May 16 was lower by 25 ± 7 per cent as compared with the June values.

These results are consistent with the hypothesis that all components of cosmic radiation of the same rigidity undergo similar reduction of flux during a Forbush decrease. In order to understand more fully the process responsible for the Forbush decrease, it will be important to measure the rigidity spectrum of protons, α -particles, and heavy nuclei and their energies over the same periods of time before and during Forbush decrease events.

Acknowledgments. This work was supported by the Office of Naval Research under contract N00019-57-1-110(19). I am grateful to Professor E. P. Ney and Dr. Phyllis S. Freier for many helpful discussions. I also wish to thank the monitoring group of the University of Minnesota for exposing the emulsions in balloon flights, and Mrs. Julianne H. Miss Barbara Berry, Miss Julia Petts, Mrs. H. Rahlenbach, Mrs. Dawn Copeland, and I. Biruta Sommer for careful scanning of emulsions.

REFERENCES

- Appa Rao, M. V. K., S. Biswas, R. R. Daniel, N. Neelakantan, and B. Peters, Abundance of heavy nuclei in the primary cosmic radiation, *Phys. Rev.*, **110**, 751-765, 1958.
- Biswas, S., P. J. Lavakare, K. A. Neelakantan, P. G. Shukla, Evidence for a Forbush type decrease in the intensity of heavy nuclei of primary cosmic radiation, *Phys. Rev.*, **118**, 591-596, 1960.
- Carmichael, H., and J. F. Steljes, Unusual cosmic ray fluctuations on July 17, and July 18, 1959, *Phys. Rev. Letters*, **3**, 392-394, 1959.
- Cester, R., A. Debenedetti, C. M. Garelli, B. Quasiati, L. Tallone, and M. Vigone, On the chemical and energy spectrum of heavy primaries in cosmic radiation, *Nuovo cimento*, **7**, 371, 1958.
- McDonald, F. B., Study of geomagnetic variations at low energies and temporal variation of the primary cosmic radiation, *Phys. Rev.*, **107**, 1386, 1953.

- Donald, F. B., and W. R. Webber, Changes in low rigidity primary cosmic radiation during the large Forbush decrease of May 12, 1959, *J. Geophys. Research*, **65**, 767-770, 1960.
- Eyer, P., The cosmic ray alpha-particle Flux during sharp Forbush intensity decreases, *J. Geophys. Research*, **65**, 3881-3887, 1960.
- Freier, E. P., J. R. Winckler, and P. S. Freier, Protons from the sun on May 12, 1959, *Phys. Rev. Letters*, **3**, 183-185, 1959.
- Gooding, C. J., Charge Distribution of multiply charged nuclei in the primary cosmic radiation, *Phil. Mag.*, **2**, 1059-1078, 1957.
- Winckler, J. R., and P. D. Bhavsar, Low energy solar cosmic rays and the geomagnetic storm of May 12, 1959, *J. Geophys. Research*, **65**, 2637-2655, 1960.
- Winckler, J. R., P. D. Bhavsar, and L. Peterson, The time variation of solar cosmic rays during July 1959 at Minneapolis, *J. Geophys. Research*, **66**, 995-1022, 1961.

(Manuscript received June 15, 1961.)



Radioactivity Produced in Discoverer XVII by November 12, 1960, Solar Protons

JOHN T. WASSON

*Geophysics Research Directorate
Air Force Cambridge Research Laboratories
Bedford, Massachusetts*

Abstract. Scintillation-spectroscopy measurements on a AgBr emulsion block from Discoverer XVII, which was flown during a period of high solar cosmic-ray activity on November 12, 1960, reveal a gamma-ray spectrum attributable to 8.4-day Ag^{106} . The disintegration rate, corrected to a probable production time of 2200 UT, November 12, 1960, is 14 dis sec^{-1} . If one assumes a (p, pn) cross section of 100 mb, and applies the thin-target formula for production of radioactivity, this corresponds to a total proton dosage of about $1.6 \times 10^8 \text{ protons cm}^{-2}$ within the emulsion, and to a value of 16 rads radiation dosage. An attempt to measure the gamma spectrum of 40-day Ag^{106} was unsuccessful, allowing the assignment of an upper limit on the disintegration rate of 1 dis sec^{-1} at the time of production. A search for 1.3-year Cd^{109} has been unsuccessful.

Introduction. The first observation of high-energy cosmic rays generated by the sun was on February 28, 1942 [Forbush, 1946]. The large solar cosmic-ray effect on November 12, 1960, was the eighth such event which has been recorded, indicating an average of one such occurrence about every 3 years. The fortuitous launching of the Discoverer XVII recoverable satellite during the peak of the solar activity has provided nuclear chemists with their first chance to apply radioactivity techniques to the measurement of properties of the solar cosmic rays. The techniques involved are much the same as those employed during recent years to investigate long- and short-term cosmic-ray intensities by studying long- and short-lived radioactivities in meteorites. The maximum in the intensity of the November 12 cosmic-ray increase as measured by neutron monitors at middle latitude stations in North America occurred at about 2000 UT [Steljes, Carmichael, and McCracken, 1961]. Discoverer XVII was launched toward the south at 1942 UT and was above the sensible atmosphere by about 2045 UT [Yagoda, 1961]. It went into a polar orbit, with an apogee of $993 \pm 4 \text{ km}$ occurring at 20°S , and a perigee at $188 \pm 3 \text{ km}$ occurring at 18°N . The period was 96.4 minutes, and the total time in orbit was 49.8 hours, after which time it was retrieved in midair over the Pacific Ocean. The next giant solar flare producing high-energy cosmic rays occurred 1 day later, on November 15, 1960.

The energy spectrum of solar cosmic rays falls rapidly with energy above a certain energy value, roughly as E^{-5} [Winckler, 1960]. The exact point where this rapid fall commences varies with the time after the flare, occurring at about 1 beV for prompt particles from the flare of February 23, 1956, and at about 200 Mev 19 hours after the same flare. Below these energies the spectrum changes much more slowly with energy. The lower limit on the energy is given by Winckler as about 40 Mev, although this is very uncertain. We shall assume that the primary protons that produced radioactivity in Discoverer XVII were in the energy range $45 < E < 400 \text{ Mev}$, where the lower limit is given by the energy of the protons necessary to penetrate the smallest amount of absorber seen by the sample.

The geomagnetic cutoff for protons of this energy occurs at about 60° geomagnetic latitude. However, the earth's magnetic field was disturbed by a major magnetic storm during the exposure period, so we shall assume a modified cutoff latitude of 50° . A calculation on this basis indicates that the satellite was bombarded by protons in the given energy range during 45 per cent of its orbit.

Since it was launched toward the south, Discoverer XVII was first above 50° latitude about 30 minutes after it was launched, or about 70 minutes after the time of the maximum cosmic-ray intensity, which occurred at about 2000 UT. According to Steljes, Carmichael, and McCracken

[1961], this maximum in the neutron-monitor measurements was the result of earth moving into a region containing magnetically trapped particles from a flare that occurred on the sun at 1320 UT. Thus, the satellite started sampling almost 8 hours after the actual occurrence of the flare on the sun.

The chemical composition of the primary components of solar cosmic particles has not been measured, but the characteristics are consistent with an assignment of a proton beam [Winckler, 1960]. The stellar abundance data of Aller [1953] gives a ratio of hydrogen to helium of about 7 : 1, which should be a lower limit for flare particles. All other elements are less abundant than hydrogen by a factor greater than 1000. This agrees well with the abundance given by Ney [1960] for galactic cosmic rays. We have assumed that all the observed radioactivity was produced by protons.

Although optical measurements on flares indicate a duration of only a few minutes, the flare particle flux at the earth lasts much longer. Neutron monitor measurements indicate a period of about 1 hour for the flux to rise and peak, at which time it falls proportional to the inverse square of the time [Winckler, 1960]. The data obtained from Geiger-counter measurements made on Explorer VII during the period November 12-15, 1960, [Van Allen, 1961] can be interpreted as fitting a t^{-2} decay following the November 12 flare.

Results and discussion. The observed radioactivity was produced in a AgBr emulsion block belonging to H. Yagoda, at AFCRL. This block was composed of Ilford G5 emulsion, measuring $10 \times 15 \times 1.3$ cm. Our measurements were on a total weight of 698 g (total weight flown was 772 g) of which 330 g were silver, and 261 g were bromine, and the remainder was gelatin. The minimum absorber path seen by the surface of the emulsion was 2.2 g cm^{-2} , and this covered a solid angle of 3 steradians. The remaining solid angle had absorbed thicknesses varying between about 5 and 25 g cm^{-2} .

Soon after the recovery of Discoverer XVII, Yagoda suggested to us that radioactivity might be detectable in his emulsion block (in which preliminary trial developments showed deep blackening from the extreme exposure). (Subsequently a development procedure was devised which suppressed the proton and alpha track

background so that the tracks of slow heavy primaries of charge $Z > 4$ could be resolved at very high magnification.) A gross gamma spectrum taken on November 21 during a 1-min count with the emulsion block sitting on the face of a 3×3 -in NaI scintillation crystal and recorded in a 100-channel analyzer is shown in curve A of Figure 1. Curve B was taken on November 22 during a 930-min count on an unexposed emulsion block. Curve C shows the differential spectrum obtained by subtracting the blank count from the count of the Discoverer XVII emulsion. The energy calibration was obtained from the K^{40} and $\text{Th C}''$ peaks, the background, and from calibration of the detection system with known gamma-ray standards.

The lower portion of Figure 1 shows a known spectrum of 8.4-day Ag^{106} , recently published by Robinson, McGowan, and Smith [1960]. As can be seen, the energies and intensities of curves A and B correspond almost exactly to those given for Ag^{106} , and we attribute the entire spectrum below 1.8 Mev to that nuclide, with the possible additional contribution of an unknown species at about 1.38 Mev. We are unable to assign the high-energy peaks.

Ag^{106} was formed from Ag^{107} by a (p,n) reaction. This type of reaction has the highest cross section of any type of reaction induced by protons of energy $45 < E < 400$ Mev incident on medium- Z targets, as indicated by data of Caretto and Wiig [1959], Ladenbauer and Vetter [1960], Fink and Wiig [1954] and Mead [1953]. The other (p,pn) products from silver and bromine isotopes are all too short-lived to have been detected at the time of our measurements, about a week after the solar flare.

In order to obtain the total counting rate from the emulsion block, we have counted a block of KCl having the same shape as the original emulsion block, and containing 157 g K. From the known isotopic abundance and half-life of K^{40} , we have calculated the efficiency of the detector with respect to 1.5-Mev gamma rays emitted in the emulsion block. The value obtained was 0.00614, or a geometrical efficiency of about 11 per cent. After multiplying the count rate by a factor of 0.81 to correct for the contribution from the Comptons from higher-energy gammas, by a factor 1.13 to allow for the absorption of the gammas in the source, and allowing for the slight differences in absorp-

d in efficiency and peak-to-total ratios of the 1.6-Mev K^{40} gamma and the 1.54-Mev gamma, obtain a value of 1.54 gammas sec^{-1} of the 1.54-Mev transition. This corresponds to a integration rate of 8.4-day Ag^{106} of 6.58 dis sec^{-1} at the time of our measurement, or 14.2 sec^{-1} at an assumed time of production of 00 UT, November 12, 1960. The number of radioactive atoms of Ag^{106} that would give this integration rate is 1.49×10^7 atoms.

According to the preceding references given for proton-induced reactions, the (p,pn) cross section peaks at about 30 Mev, and then levels off to a constant value for higher energies. The measured values vary from about 60 to 600 mb. The present case is complicated by the nuclear isomerism of Ag^{106} . The 8.4-day isomer has a $6+\pi$, whereas the 24-min isomer has a $1+\pi$ spin. It is uncertain which is the ground state. However, the data of *Meadows, Diamond, and Sharp* [56], and Matsuo (private communication, 1961) indicate that in this energy range the high-spin state should be formed in a higher abundance than the low-spin state. We shall assume a cross section, σ_{106} , for the production of 8.4-day Ag^{106} of 100 mb, which is independent of energy.

From the thin-target formula,

$$D_p = \frac{N_{106}}{N_{107}\sigma_{106}}$$

for given values of σ_{106} , N_{107} (the number of Ag^{107} atoms), and N_{106} (the number of Ag^{106} atoms at the time of production), we can calculate D_p , the proton dosage within the emulsion block during the flare. The value thus obtained is 1.6×10^3 protons cm^{-2} . If we assume an isotropic flux and that all protons entered through the 1.6-Mev 'window,' the dosage in space was $1.6 \times 10^3 \text{ cm}^{-2}$. Furthermore, since the target is clearly not thin (the smallest target thickness, 4.5 g cm^{-2} is just the range of a 50-Mev proton in a AgBr medium), the stated value is only a lower limit. In an effort to correct for the attenuation of the time during which the satellite was below the geomagnetic cutoff latitude, we shall assume that it sampled 45 per cent of the total flux, corresponding to a cutoff latitude of 50° . The integrated flux is then 1.5×10^3 protons cm^{-2} . It should be possible to obtain a value for the instantaneous flux from our value of the inte-

grated flux. *Winckler* [1960], gives the formula,

$$I = I_0 \frac{t_0^2}{t}$$

where t_0 is a constant equal to 1 hour, and corresponds to the beginning of the t^{-2} decay of the flux. I_0 is the flux at t_0 , and I is the flux at a time $t > t_0$. If our data are fitted to the curve obtained by assuming that Discoverer XVII started sampling 7.5 hours after the flare, an initial value for the instantaneous flux of 6.4×10^4 protons sec^{-1} is obtained. *Van Allen* [1961] reports fluxes measured by Explorer VII at 2330 UT on November 12 and at 0356 UT on November 13 of 2.1×10^4 and 1.1×10^4 particles $\text{cm}^{-2} \text{ sec}^{-1}$. Our calculated fluxes at these times are 3.5×10^4 and 1.7×10^4 protons $\text{cm}^{-2} \text{ sec}^{-1}$, in satisfactory agreement with Van Allen's results.

It is possible to calculate the radiation dosage from the value for the proton dosage within the emulsion (1.6×10^3 protons cm^{-2}). If the approximate energy of the protons entering the emulsion is 100 Mev, and they pass through an approximate absorber thickness of 8.0 g cm^{-2} , then one can calculate from range-energy relationships that the average energy loss is 50 Mev per proton. The approximate area of the emulsion having an 8 g cm^{-2} thickness is 88 cm^2 . From these numbers, plus the weight of the emulsion, 698 g, the energy dosage per gram of emulsion can be calculated. The value obtained is 1.6×10^3 ergs g^{-1} , corresponding to a radiation dosage of 16 rads. This compares with a value of about 50 rads obtained by *Yagoda* [1961] from grain-density measurements on the same emulsion. It should be noted, however, that Yagoda has measured the total radiation dosage due to all types of radiations, whereas we have only measured that fraction due to protons within the given energy limits.

Owing to the unfortunate interpretation of our original data (top curves in Figure 1) as indicating essentially no radioactivity, no attempts were made to find additional activities until mid-January 1961, some 2 months after the flare. Possible activities still measurable at that date included 40-day Ag^{105} , 120-day Se^{75} , and 1.3-yr Cd^{109} . Since the emulsion block had now been developed, we isolated the silver from the hypo solution and looked for Ag^{105} . An upper limit on the number of atoms of Ag^{105} on Novem-

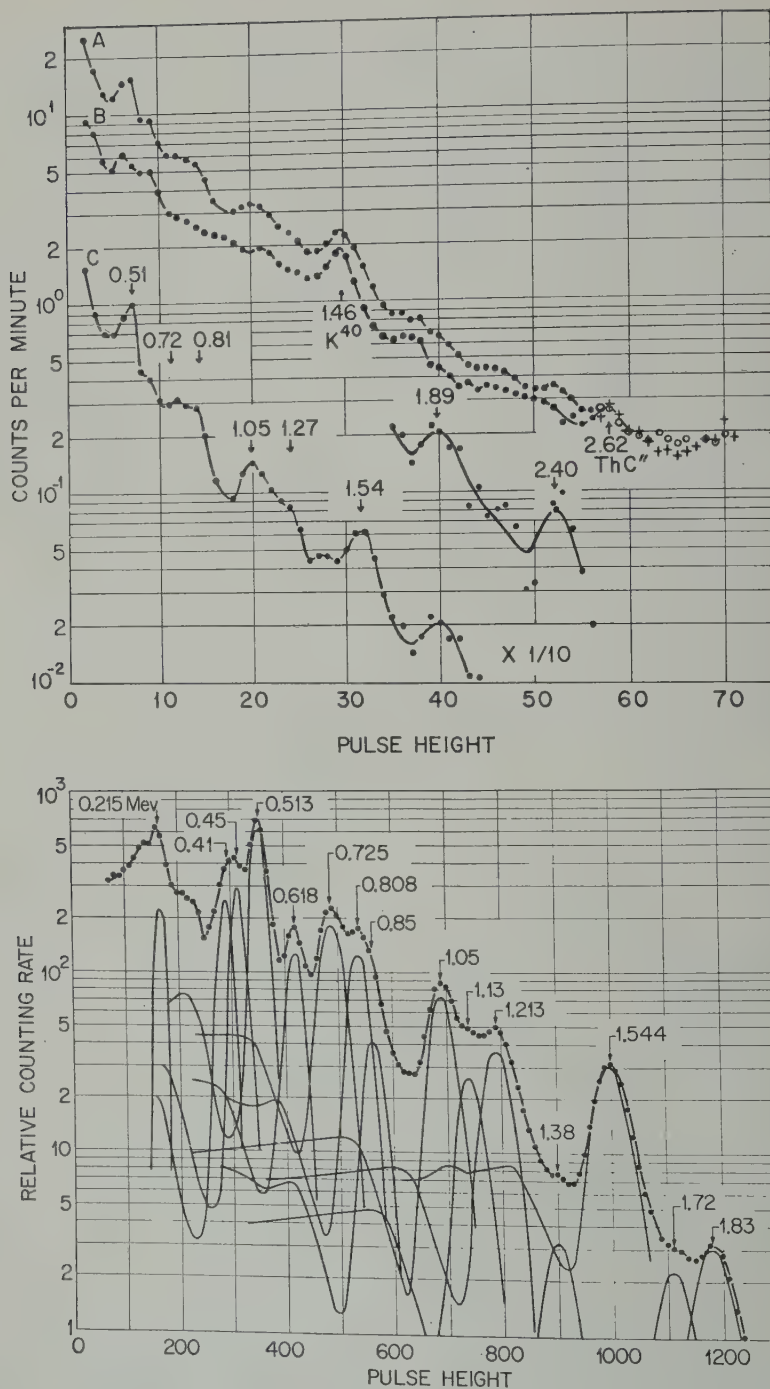


Fig. 1. The upper portion of this figure shows: A. The gross gamma spectrum of $15 \times 10 \times 1$ cm AgBr emulsion block from Discoverer XVII, the block resting on the face of 3×3 -in NaI detector. B. Spectrum obtained by measuring an unflown AgBr emulsion block of identical dimensions. C. Spectrum obtained by subtracting curve B from Curve A, and interpreted as ^{106}Ag . The lower portion of the figure shows a gross spectrum of ^{106}Ag given by Robinson, McGowan and Smith [1960].

12, 1960, of 5×10^6 was established. On the basis that most of the radioactivity was still on the developed plates, a new count was made about 20 per cent of the original sample. Identifiable activity was observed. A search for Cd^{109} from the hypo solution that involved use of a low-background X-ray counter (Kalkstein, unpublished material) was also unsuccessful.

The method of studying radioactivity produced by cosmic rays offers great promise for the measurement of high-intensity particle spectra, as those produced in the Van Allen belts during solar flares. If one takes advantage of the variation with energy of the ratio of the cross sections for various reactions, as (p, pn) and $(p, p2n)$, then not only the flux but also the energy spectrum of these particle streams should be measurable by such techniques.

Acknowledgments. I would like to express my thanks to Dr. H. Yagoda for the sample, and to my colleagues, Dr. P. J. Drevinsky, Dr. I. Kalkstein, and Mr. K. Shivanandan, for their helpful discussions. I am also grateful to Anahid Thomasian and Mr. Edward C. Goble for performing chemical separations, and to Mr. R. L. Robinson for the Ag^{108} spectrum. This work was supported in part by the U. S. Atomic Energy Commission.

REFERENCES

Lawrence H., *The Atmospheres of the Sun and Stars*, Ronald Press Company, New York, 1953.
 Lingenfelter, Albert A., Jr., and Edwin O. Wiig, Interaction of yttrium with protons of energy between 60 and 240 Mev, *Phys. Rev.*, **115**, 1238, 1959.

- Fink, Richard W., and Edwin O. Wiig, Reactions of cesium with protons at 60, 80, 100, 150, and 240 Mev, *Phys. Rev.*, **96**, 185, 1954.
 Forbush, Scott E., Three unusual cosmic-ray increases possibly due to charged particles from the sun, *Phys. Rev.*, **70**, 771, 1946.
 Ladenbauer, Inge-Maria, and Winsberg, Lester, Interaction of high-energy protons and alpha particles with iodine-127, *Phys. Rev.*, **119**, 1368, 1960.
 Meadows, J. W., Excitation functions for proton-induced reactions with copper, *Phys. Rev.*, **91**, 885, 1953.
 Meadows, J. W., R. M. Diamond, and R. A. Sharp, Excitation functions and yield ratios for the isomeric pairs $\text{Br}^{80,80m}$, $\text{Co}^{58,58m}$, and $\text{Sc}^{44,44m}$ formed in (p, pn) reactions, *Phys. Rev.*, **102**, 190, 1956.
 Ney, E. P., Cosmic rays as received at the earth, *Astrophys. J. Suppl.*, **4**, 371, 1960.
 Robinson, R. L., F. K. McGowan, and W. G. Smith, Decays of Rh^{106} and Ag^{106} , *Phys. Rev.*, **119**, 1692, 1960.
 Steljes, J. F., H. Carmichael, and K. F. McCracken, Characteristics and fine structure of the large cosmic-ray fluctuations in November 1960, *J. Geophys. Research*, **66**, 1363-1377, 1961.
 Van Allen, J. A., Explorer VII observations of solar cosmic rays, November 12-23, 1960, *Bull. Am. Phys. Soc.* (II), **6**, 276, 1961.
 Winckler, J. R., Primary cosmic rays, and discussion, reprinted in *NASA Technical Note, D-588*, December 1960.
 Yagoda, Herman, Radiation studies from nuclear emulsions and metallic components recovered from polar satellite orbits, *Geophysics Research Directorate Research Note No. 54*, to be published in report of COSPAR meeting, Florence, April 10-14, 1961.

(Manuscript received May 9, 1961;
 revised June 6, 1961.)



The Lunar Neutron Flux

R. E. LINGENFELTER, E. H. CANFIELD, AND W. N. HESS

*Lawrence Radiation Laboratory, University of California
Livermore, California*

Abstract. The equilibrium neutron-leakage spectrum at the lunar surface and the neutron capture as a function of depth beneath the surface have been calculated for a number of assumed lunar surface compositions. The leakage spectrum is found to be particularly sensitive to hydrogen abundance; an experiment is accordingly proposed to determine the hydrogen content of the lunar surface by neutron flux measurements near the moon.

INTRODUCTION

With the advent of vehicles that can come to the moon, or even land on it, much attention has been given to the problem of determining the chemical composition of the moon, and, in particular, of its surface. The moon is of interest not only for its own sake but because the lunar surface, undisturbed as it is by atmospheric processes such as erosion and chemical weathering, should hold many clues to both its own origin and that of the earth and planets.

Considerable thought has been given to detecting characteristic X rays from the lunar surface, either those that exist naturally (Arnold, private communication) or X rays generated artificially by bombardment of the surface (Anderson and Martina, 1960). Much information about the composition of the top few feet can be obtained this way.

Neutrons can also be used to study the composition of the moon. The interaction of primary cosmic rays with the material of the lunar surface generates a high-energy neutron flux at the surface. The calculated equilibrium energy spectrum and spatial distribution of this neutron flux are strongly dependent on the assumed composition of the lunar surface and are particularly sensitive to the hydrogen content. In the present study, multigroup diffusion calculations of the equilibrium flux are made for a number of assumed lunar surface compositions. The surface neutron-leakage current and the neutron absorptions as a function of depth and material are presented. A possible experiment for the determination of the hydrogen content of the lunar surface from neutron flux

measurements near the moon is also described.

DIFFUSION CALCULATIONS

The diffusion equation for the equilibrium scalar neutron flux $\phi(E, r)$ may be written:

$$-\nabla \cdot D(E, r) \nabla \phi(E, r) + \Sigma(E, r) \phi(E, r) - \int_0^\infty dE' \Sigma(E' \rightarrow E, r) \phi(E', r) = S(E, r)$$

where

$$\begin{aligned} \phi(E, r) &= n(E, r) v. \\ n(E, r) dE &= \text{number of neutrons/cc in the energy interval } E, E + dE. \\ \Sigma(E, r) &= \text{total macroscopic cross section.} \\ \Sigma(E' \rightarrow E, r) &= \text{macroscopic cross section for scattering from } dE' \rightarrow dE, \\ D(E, r) &= \text{diffusion coefficient} = 1/[3\Sigma_{tr}(E, r)]. \\ \Sigma_{tr}(E, r) &= \text{macroscopic transport cross section.} \\ S(E, r) &= \text{neutron source} = \text{number of neutrons/cc-sec.} \end{aligned}$$

Multiplying by dE and integrating from E_{i-1} to E_i gives the usual multigroup diffusion equation for the i th energy group ($E_{i-1} < E < E_i$):

$$-\nabla \cdot D^i(r) \nabla \phi^i(r) + \Sigma^i(r) \phi^i(r) - (\text{Sum})_i \Sigma^{i,j}(r) \phi^j(r) = S^i(r)$$

This equation is solved by the Zoom code [Stuart, Canfield, Dougherty, and Stone, 1958], developed for the IBM 709 at the Lawrence Radiation Laboratory at Livermore. In the present calculations the moon was represented as a semi-infinite medium, of density 1 g/cc.

The cross sections are essentially those from BNL 325 [Hughes and Schwartz, 1958], weighted with a Maxwellian thermal-neutron spectrum and a $1/E$ tail. The transfer coefficients for scattering between energy groups are based on the assumption that the lunar materials may be represented as Maxwell gas moderators which scatter isotropically in the center of mass. These coefficients were calculated with Sophist I, another IBM-709 code developed by LRL [Canfield, Collins, Feris, and Stuart, 1960]. Above 1 Mev, the effects of anisotropy in the elastic scattering were compensated for by corrections linear in the average value of the cosine of the scattering angle.

Neutron source. The magnitude, energy, and spatial distribution of neutrons generated by cosmic rays in the lunar material can be inferred from knowledge of neutrons generated by cosmic rays in the earth's atmosphere [Hess, Canfield, and Lingenfelter, 1961]. But the production rate of neutrons at the lunar surface is higher than in the earth's atmosphere for three reasons: the higher average cosmic-ray intensity at the lunar surface, the higher average atomic mass of lunar material, and the added neutron production due to π mesons.

The cosmic-ray intensity at the surface of the moon is larger than in the upper atmosphere of the earth because the magnetic field of the moon is very small. Thus there should be no energy cutoff as a function of latitude in the cosmic-ray flux incident on the lunar surface. The cosmic-ray intensity at the lunar surface should be the same as that above the poles of the earth, or about 2.0 times the earth's global average.

The production cross section for evaporation neutrons was founded by Graves and Rosen [1953] and by Gross [1956] to be roughly proportionate to \sqrt{A} (A = atomic no.). A lunar composition similar to that of chondritic meteorites (to be discussed later) would have an average mass of $A = 24.5$, as against $A = 14.5$ for the earth's atmosphere. This gives an average neutron-production cross section for the lunar material 1.3 times that for the earth's atmosphere, or a source strength of 12 ± 2 neutrons/cc-sec from evaporation and knock-on processes.

Pi mesons are made in cosmic-ray interactions with matter. In the earth's diffuse atmosphere, essentially all these decay with a half-life of 2×10^{-8} sec before capture. However,

in the more dense lunar surface, essentially every π^- meson will be captured and will release about 4 neutrons (Kaplan, 1960, private communication); the π^+ mesons are not captured by nuclei because of Coulomb repulsion. Cosmic-ray accelerator data show that about one π^- meson is produced per incident proton [Puppi and Dallaporta, 1952; Johnson, 1955]. Assuming an incident proton flux at the lunar surface to be the same as at the earth's poles, about 2 protons/cc-sec [Puppi and Dallaporta, 1952], we have 8 ± 3 neutrons/cc-sec from (π^-, n) events.

Adding the evaporation, knock-on, and π meson components, we then assume the total neutron strength for the moon to be 20 ± 3 neutrons/cc-sec, as opposed to 4.6 neutrons/cc-sec for the global average of the earth's atmosphere [Hess, Canfield, and Lingenfelter, 1961].

For those neutrons not resulting from (π^-, n) events, the energy distribution of the source is assumed to be the same as that for neutrons made in the earth's atmosphere; that is, 20 per cent resulting from knock-on processes, with a spectrum

$$N(E) dE = kE^{-2} \exp(-160E^{-2}) dE$$

and 80 per cent resulting from evaporation processes, with a spectrum

$$N(E) dE = kE \exp(-E/\theta) dE$$

with $\theta = 1$ Mev. The neutrons resulting from (π^-, n) processes are assumed to have the same evaporation spectrum. The combined neutron source is, therefore, 88 per cent evaporation spectrum and 12 per cent knock-on spectrum.

Almost all the evaporation-spectrum neutrons lie below 10 Mev, while those of the knock-on spectrum lie above. However, the diffusion calculation treated only those neutrons below 10 Mev; the source spectrum for the calculation was therefore modified to include those neutrons born at higher energies, that were scattered below 10 Mev before being captured or leaking. This was done by a simple conservation argument similar to that used for the neutron calculations in the earth's atmosphere [Hess, Canfield, and Lingenfelter, 1961], which indicates that 48 per cent of those neutrons born above 10 Mev are captured or leak before scattering down to below that energy. Subtracting this percentage from the total source, we com-

TABLE 1. Composition of Chondritic Meteorites

Element	Abundance, atoms/gram	Thermal-Capture Cross Section, barns/gram
	0.0143	0.0
	0.000207	0.000105
	0.00387	0.000244
	0.000321	0.000074
	0.00414	0.000663
	0.000019	0.000004
	0.000431	0.000224
	0.000024	0.000050
	0.000255	0.000102
	0.000009	0.000053
	0.000031	0.000098
	0.000023	0.000306
	0.00294	0.007431
	0.000010	0.000352
	0.000149	0.000716
	0.87×10^{-6}	0.000003
	2.0×10^{-6}	0.000002
	0.28×10^{-6}	0.000003
	0.37×10^{-8}	0.000009
	0.27×10^{-8}	0.000015
	0.08×10^{-6}	0.000003
	0.28×10^{-8}	0.000130
	0.23×10^{-8}	0.000002
	0.41×10^{-8}	0.000002
Total		0.010591

effective source strength for the diffusion equation to be 18.8 neutrons/cc-sec.

The spatial distribution of the neutron source is taken as exponentially decreasing from the surface with the form $e^{-d/155}$, where d is the depth in grams beneath the lunar surface. The diffusion thickness of $155 \pm 10 \text{ g}^{-1}$ was found experimentally (Patterson, 1960, private communication) by bombarding concrete with thermal neutrons. This is the same e -folding thickness found by Hess, Patterson, Wallace, and Urey [1959] for the neutrons in the earth's atmosphere.

Lunar Composition. A number of possibilities have been proposed for the composition of the lunar surface [Urey, 1959; 1960]. In this paper we consider compositions similar to that of chondritic meteorites, with variations in the thermal-capture cross sections and in the hydrogen content, and compositions similar to those of granite, basalt, and the tektites. Ninninger [1943] stated that the tektites may in fact have originated from the moon.

The composition of the chondrites (chondritic

meteorites) is given by Urey and Craig [1953]; the abundance of elements above $Z = 29$ is given by Suess and Urey [1956]. This composition is shown in Table 1, together with the thermal-capture cross section per gram. The table gives the composition for the weighted average of the high- and low-iron groups of the chondrites; the low-iron group being neutronically equivalent to a reduction of the total $1/v$ capture cross section by 12 per cent and the high-iron groups equivalent to an increase of 16 per cent. Urey [1959] has calculated that if the average composition of the moon is similar to that of the chondrites, differing only in the abundance of iron, then an iron-to-silicon-atom ratio of about 0.35 to 1 is required to give the observed density. This is neutronically equivalent to a reduction of 35 per cent in the total $1/v$ capture cross section.

The second important variation in composition affecting the equilibrium neutron spectrum is the hydrogen content. In the chondrites, water abundances of about 0.25 per cent (or a H/Si atomic ratio of 0.04) have been found [Urey, 1959]. Urey also points out that if the water content of chondritic material were as high as 2 per cent (or a H/Si atomic ratio of 0.32), then its density would also be the same as the average lunar density. Of course, whether the surface material is representative of the average composition depends on the amount of differentiation that the lunar surface has undergone, and also on the depth of meteoric dust, but most probably the hydrogen content of the surface material should approximate that of the average composition within a reasonable range.

In addition, equilibrium spectrum calculations have been made assuming compositions similar to granite, basalt, and the tektites. These compositions, given by Rankama and Sahama [1950] are shown in Table 2, together with the thermal-capture cross section per gram. These calculations were made at the average lunar surface temperature of -9°C , whereas in fact the lunar surface is usually at either the 'day' or 'night' extremes of $+134^\circ\text{C}$ or -153°C , i.e., the night-to-day variation is 2.5 times the minimum absolute temperature. But the effect of this variation on the neutron-leakage spectrum is only slight (a variation in the flux of the order of 10 per cent below 0.1 ev), and the effect on the rate of capture versus depth is insignificant.

TABLE 2. Composition of Granite, Basalt, and the Tektites

Granite			Basalt			Tektites		
Element	Abundance atoms/gram	Thermal-Capture Cross Section barns/gram	Element	Abundance atoms/gram	Thermal-Capture Cross Section barns/gram	Element	Abundance atoms/gram	Thermal-Capture Cross Section barns/gram
O	0.0183	0.	O	0.0169	0.	O	0.0184	0.
Si	0.00704	0.00113	Si	0.00492	0.00079	Si	0.00775	0.00113
Al	0.00171	0.00039	Al	0.00186	0.00043	Al	0.00131	0.00039
Fe	0.000267	0.00068	Fe	0.000937	0.00237	Fe	0.000246	0.00068
Mn	0.000010	0.00013	Mn	0.000026	0.00034	Mn	0.000009	0.00013
Mg	0.000131	0.00001	Mg	0.000918	0.00006	Mg	0.000147	0.00001
Ca	0.000214	0.00009	Ca	0.000962	0.00042	Ca	0.000292	0.00009
Na	0.000676	0.00034	Na	0.000604	0.00031	Na	0.000088	0.00034
K	0.000526	0.00109	K	0.000194	0.00040	K	0.000318	0.00109
H	0.000562	0.00019	H	0.001084	0.00036			
Ti	0.000029	0.00017	Ti	0.000101	0.00059			
P	0.000016	0.	P	0.000038	0.00001			
Total 0.00422			Total 0.00608			Total 0.00422		

NEUTRON EQUILIBRIUM SPECTRUM

The calculated neutron equilibrium leakage spectrum for the lunar surface is shown in Figures 1 and 2 for a number of assumed compositions. Figure 1 shows the leakage spectrum for chondritic material, chondritic material with a 35 per cent decrease in the total $1/v$ capture cross section, a 10 per cent increase, and a 50 per cent increase, and chondritic material with hydrogen added in a H/Si atomic ratio of 0.1 and 1.0. It can be seen that a variation of the $1/v$ capture cross section affects only the low-energy leakage and that the high-energy leakage is unchanged; whereas the addition of hydrogen reduces the amount of high-energy leakage and increases the low-energy leakage. This effect provides a possible means of measuring the hydrogen content of the lunar surface material, as will be discussed in the following section. Figure 2 shows the leakage spectrum for chondrites, basalt, granite, and tektites.

As Figures 1 and 2 indicate, the integrated leakage is quite sensitive to the hydrogen content. For chondritic material, 35.9 per cent of the neutrons leak into space. This leakage decreases to 30.3 per cent with the addition of 0.1 hydrogen atom per silicon atom, and to 17.3 per cent with 1.0 hydrogen per silicon. At the same time, the integrated leakage is quite insensitive to changes in the $1/v$ capture cross section, a 35 per cent decrease in capture cross section yielding only a

0.6 per cent increase in the leakage, and a 50 per cent increase yielding only a 0.7 per cent decrease. Basalt, granite, and tektites have leakages of 29.7, 32.3, and 40.0 per cent respectively.

The total neutron capture as a function of depth beneath the lunar surface is shown in Figures 3 and 4 for the same compositions. Figure 3 shows the variation of total capture as a function of depth for chondritic material with H/Si atomic ratios of 0, 0.1, and 1.0. The spatial dependence of the capture is insensitive to ± 50 per cent changes in the $1/v$ capture cross section. Figure 4 shows the spatial dependence for chondrites, basalt, granite, and tektites. Particular $1/v$ reaction rates are in the ratio of their thermal-capture cross section per gram to the total thermal-capture cross section per gram (both given in Tables 1 and 2).

A POSSIBLE EXPERIMENT

As a result of these calculations on the neutron equilibrium leakage spectrum, it appears that a simple experiment might be performed, using the neutron spectrum at or near the lunar surface, to determine the hydrogen content of the lunar surface material. There are three types of neutron detectors, having different energy sensitivities, that might be used for this experiment: (a) a detector with an energy sensitivity varying as $1/v$, e.g., a $B^{10}F_3$ proportional counter.

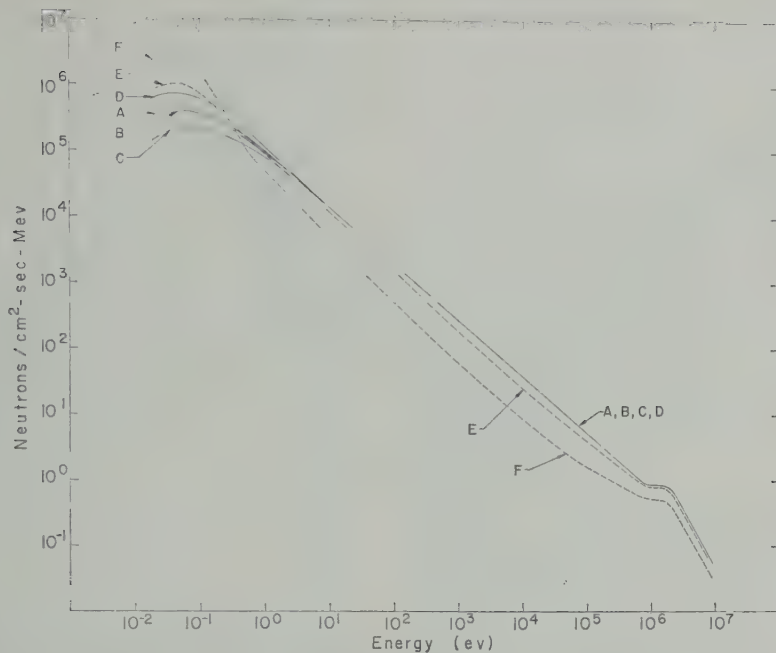


Fig. 1. The calculated neutron equilibrium leakage spectrum for the lunar surface for compositions *A*, chondritic material; *B*, chondritic material with a 10 per cent increase in the total $1/v$ capture cross section; *C*, with a 50 per cent increase in $1/v$ capture; *D*, with a 35 per cent decrease in $1/v$ capture; *E*, chondritic material with 0.1 H/Si atom; and *F*, with 1.0 H/Si atom.

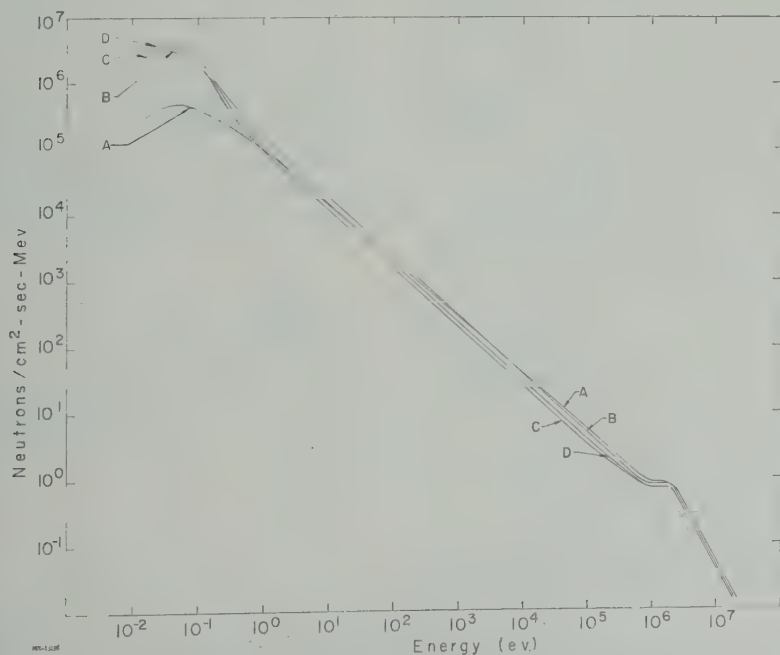


Fig. 2. The calculated neutron equilibrium leakage spectrum for the lunar surface for compositions *A*, chondrites; *B*, tektites; *C*, basalt; and *D*, granite.

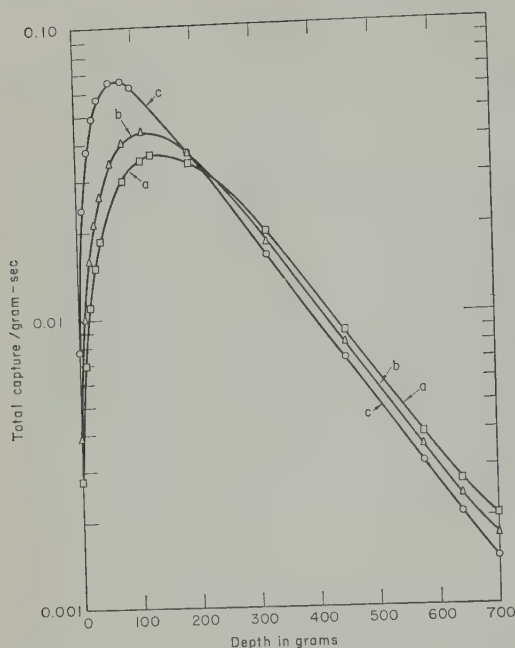


Fig. 3. The total neutron capture vs. depth beneath the lunar surface for compositions A, chondritic material; B, chondritic material with 0.1 H/Si atom; and C, with 1.0 H/Si atom.

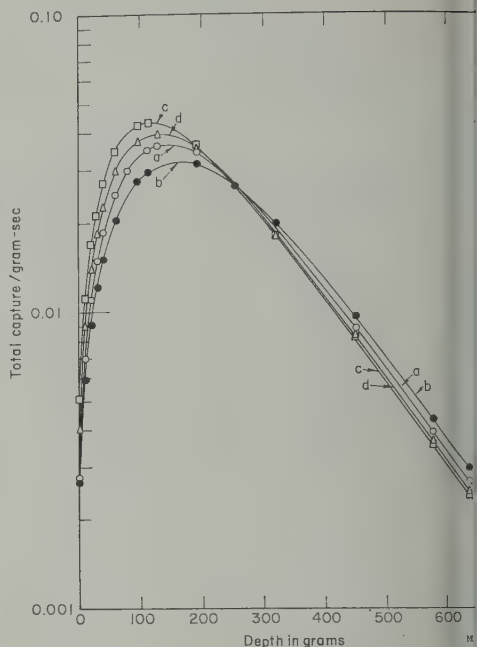


Fig. 4. The total neutron capture vs. depth beneath the lunar surface for compositions: (a) chondrites; (b) tektites; (c) basalt; and (d) granite.

(b) a counter with a flat response at higher energy, e.g., a $1/v$ counter covered with roughly 1.5 inches of a moderating material such as Epon. This type has an energy response which is nearly flat from a few kev to about 10 Mev; (c) a counter with a $1/v$ response only above some cutoff energy, e.g., a cadmium-covered $1/v$ counter which is sensitive only above about 1/2 ev. Some of these detectors have already been flown successfully to measure the neutron flux in space near the earth.

Owing to the uncertainty in the absolute source strength and the total thermal-capture cross section, a ratio of at least two detector measurements at different energies would be necessary to make a determination of the hydrogen content. Although three detectors would improve the accuracy and provide a check on internal consistency, two detectors could give very reasonable accuracy and are more acceptable from the standpoint of weight and power requirements.

The most sensitive ratio is that of detector types *a* and *b*, and Figure 5 shows curves of the ratio

$$R = \frac{\text{count rate for detector } b}{\text{count rate for detector } a}$$

for variations in the $1/v$ capture cross section vs. the H/Si atomic ratio. From this curve it can be seen that for a particular value of the ratio R a very reasonable estimate can be made of the hydrogen-to-silicon ratio. For example, the two variations in chondritic composition

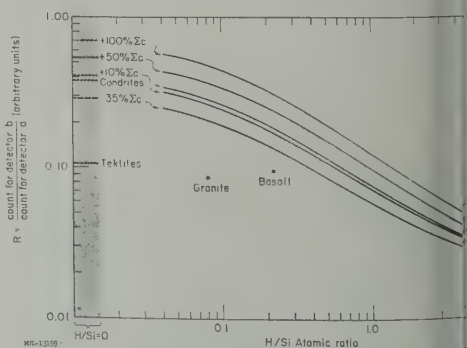


Fig. 5. The ratio $R = \frac{\text{count rate for detector } b}{\text{count rate for detector } a}$ vs. H/Si atomic ratio for variations in the total capture cross section of chondritic material.

the same density as the average lunar (a decrease of 35 per cent in $1/v$ capture) addition of 0.32 hydrogen atom per give a factor-of-2 difference in the ratio, R , and would be quite easily distinguishable. In general, the value of the ratio can be determined to nearly the same accuracy as the uncertainty in the total $1/v$ cross section (for values of H/Si greater than 0.1). Typically, if the measured value were 0.2, a ± 20 per cent uncertainty in the cross section would give roughly a ± 30 per cent uncertainty in the H/Si ratio.

The most important abundance to ascertain in determining the capture cross section is that of iron. However, iron is one of the materials that can be determined quite easily by X-ray fluorescence (Arnold, 1960, private communication). In fact, both neutron and X-ray measurements could be made at the same stage of exploration.

Another point is the possible masking of the lunar material by a covering of meteoritic rubble. In this event diffusion calculations indicate that a lunar material of H/Si = 0.32 would still be detectable for a covering of less than 25 grams/cc.

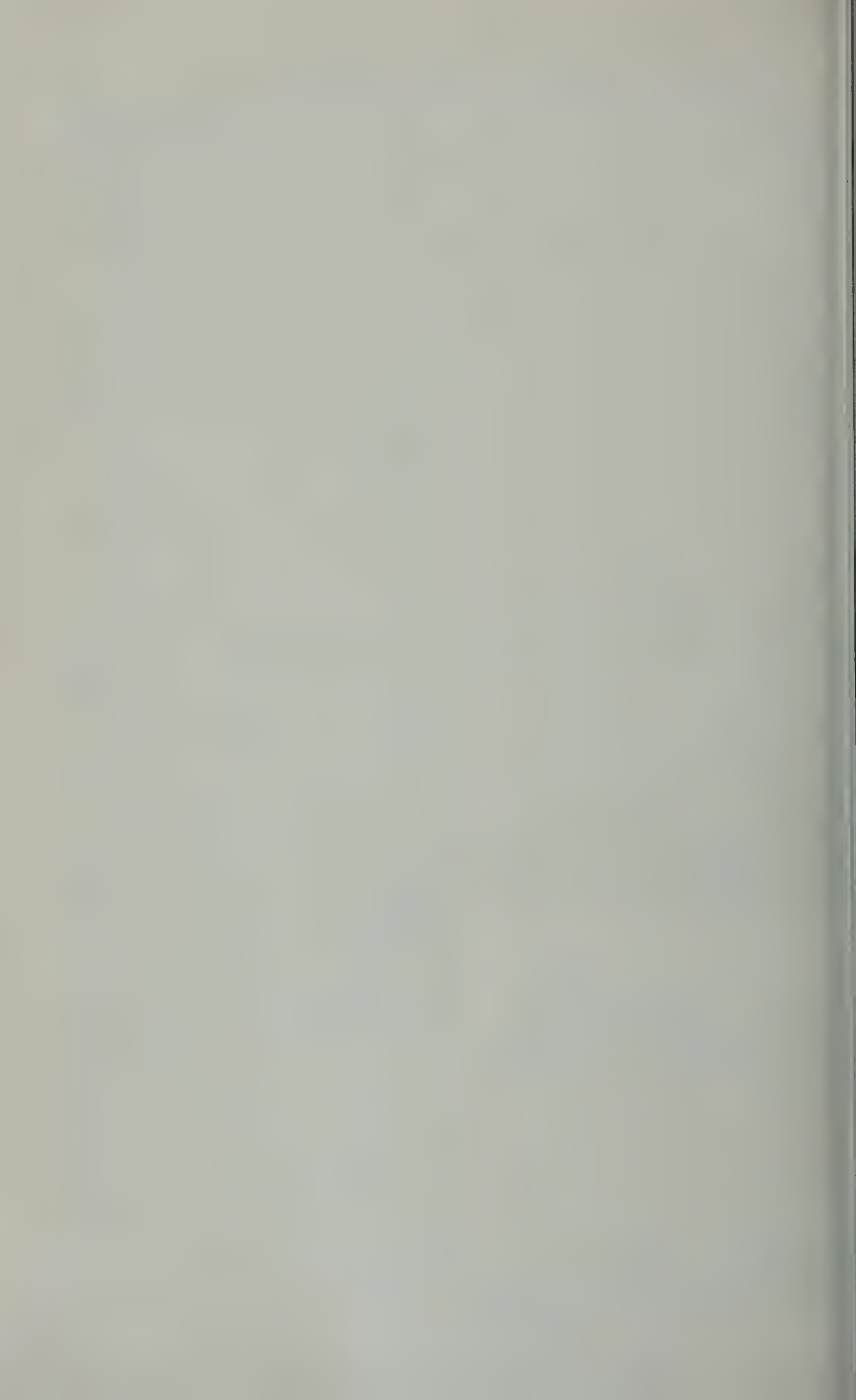
It appears that a relatively easy experiment can be performed to gage the hydrogen content of the surface of the moon by measuring neutron flux in the vicinity of the moon.

Acknowledgments. We wish to thank Professor H. C. Urey of the University of California, for his valuable discussion and suggestions regarding the possible composition of the lunar surface. This work was done under the auspices of the Atomic Energy Commission.

REFERENCES

- Arnold, E. H., W. H. Collins, R. P. Freis, and R. Stuart, 9 Sophist I, *Lawrence Radiation Lab. Report, UCRL-5956* (OTS), 1960.
- Chupp, E. L., and L. Rosen, Distribution in energy of neutrons from the interaction of 14-Mev neutrons with some elements, *Phys. Rev.*, **89**, 343, 1953.
- Gross, E., The absolute yield of low-energy neutrons from 190-Mev proton bombardment of gold, silver, nickel, aluminum, and carbon, *Lawrence Radiation Lab., (Berkeley, Calif.) Rept. UCRL-3330*, Feb. 29, 1956 (OTS).
- Hess, W. N., E. H. Canfield, and R. E. Lingenfelter, Cosmic-ray neutron demography, *J. Geophys. Research*, **66**, 665-677, 1961.
- Hess, W. N., H. W. Patterson, R. Wallace, and E. L. Chupp, Cosmic-ray neutron energy spectrum, *Phys. Rev.*, **116**, 445, 1959.
- Hughes, D. J., and R. B. Schwartz, Neutron cross sections, 2nd ed., *Brookhaven National Lab., Rept. BNL-325*, July 1, 1958 (OTS).
- Johnson, W. R., Multiple meson production in emulsions exposed to the bevatron beam, *Lawrence Radiation Lab., (Berkeley, Calif.) Rept. UCRL-2979*, May 1955 (OTS).
- Nininger, H. H., *Sky and Telescope*, **2**, 12, February 1943, and 8, March 1943.
- Puppi, G., and N. Dallaporta, The equilibrium of the cosmic ray beam in the atmosphere, in *Progress in Cosmic Ray Physics*, **1**, North-Holland Publishing Company, Amsterdam, 317-394, 1952.
- Rankama, K., and Th. G. Sahama, *Geochemistry*, University of Chicago Press, Chicago, pp. 31, 159, 166, 1950.
- Schrader, C. D., and E. F. Martina, Analysis of the moon's surface by nuclear reactions, *Lawrence Radiation Lab. (Livermore, Calif.) Rept. UCRL-5916*, April 5, 1960 (OTS).
- Stuart, R. N., E. H. Canfield, E. E. Dougherty, and S. P. Stone, Zoom: a one-dimensional, multi-group neutron diffusion theory reactor code for the IBM-704, *Lawrence Radiation Lab. (Livermore, Calif.) Rept. UCRL-5293*, November 19, 1958 (OTS).
- Suess, H. E., and H. C. Urey, Abundance of the elements, *Revs. Mod. Phys.*, **28**, 53-74, 1956.
- Urey, H. C., Preliminary revised chapters for *The Planets*, Yale University Press, October 15, 1959.
- Urey, H. C., Lines of evidence in regard to the composition of the moon, in *Space Research, Proceedings of the First International Space Science Symposium, Nice, January 11-16, 1960*, North-Holland Publishing Company, Amsterdam, 1114-1121, 1960.
- Urey, H. C., and H. Craig, The composition of the stone meteorites and the origin of meteorites, *Geochim. et Cosmochim. Acta*, **4**, 36-82, 1953.

(Manuscript received June 8, 1961.)



Transresonant Electron Acceleration

E. N. PARKER

*Enrico Fermi Institute for Nuclear Studies and Department of Physics
University of Chicago, Chicago 37, Illinois*

Abstract. Helliwell and Bell have suggested synchronous acceleration of electrons by electromagnetic waves in the whistler range in the geomagnetic field. The acceleration of trapped electrons by electromagnetic waves is here generalized to include nonsynchronous or transresonant acceleration. It is shown that whistlers will scatter the electron velocity, by an amount inversely proportional to the square root of the time rate of change of the whistler frequency, during the time that the whistler frequency is passing over the electron gyrofrequency. Only electrons with initial energy above about 10^2 ev can see the Doppler-shifted whistler frequency at the gyrofrequency and be affected by this transresonant acceleration. Thermal electrons experience no effect. Current observation of whistlers in the outer geomagnetic field is too sketchy to allow a determination of transresonant acceleration rates, so that the relative importance, as compared, for example, with the strong-whistler synchronous acceleration of Helliwell and Bell, cannot be assessed. The importance of whistler acceleration can be established when the whistler frequency spectrum and recurrence rate are known throughout the geomagnetic field.

Introduction. Helliwell and Bell [1960] have pointed out recently that the naturally occurring whistler radiation may accelerate relativistic Van Allen electrons in a manner analogous to a deuterocyclotron. Whistlers exhibit a generally increasing frequency, and the idea is that with sufficiently strong whistler the relativistic mass increase of the accelerating electron would only decrease the electron gyrofrequency, keeping the electron in phase with the descending whistler frequency. They point out the similarity of phase stability in the laboratory synchrotron. The process would apply to any electron whose velocity along the geomagnetic field was sufficient to Doppler-shift the wave frequency to the electron gyrofrequency.

The conditions for synchronous acceleration are readily demonstrated. An electron with a velocity v exactly in phase with an oscillatory electric field of amplitude E can gain energy at a rate no greater than eEv , which must be sufficient to decrease the electron gyrofrequency sufficiently as the whistler frequency is decreasing. The rate of mass increase due to E is greater than eEv/c^2 . If the relativistic electron mass is m' , then the fractional rate of change of the electron cyclotron frequency Ω is $d\Omega/dt = -eEv/m'c^2$. The characteristic time for decreasing Ω is the reciprocal, $\tau = m'c^2/eEv$, and must be no more than the charac-

teristic time of change of frequency of a whistler, i.e., of the general order of magnitude of 1 second or less. A generous lower limit on τ is obtained with the minimum value of $m'c^2$, viz., 500 kev or 0.8×10^{-6} erg, and the maximum value of v , viz., $v = c$. We find that this optimistic lower limit gives $\tau = 1.6$ seconds for fields as strong (10^{-3} volt/m = 3×10^{-8} statvolt/cm) as the concentrated ones that Helliwell and Bell propose to produce artificially. Even this optimistic estimate of τ is too large for most whistlers. Unless the whistler amplitude is for some reason very much stronger than 10^{-3} volt/m, it is insufficient to produce synchronism; both the changing field strength, along a geomagnetic line of force, and the changing Doppler-shift of the particle accelerating back and forth in the geomagnetic mirror, change the electron gyrofrequency at a much greater rate. Altogether it does not seem that synchronism and phase stability are the point of prime interest, unless whistler fields are at least 10^{-2} volt/m.

On the other hand, some kind of acceleration of Van Allen electrons appears necessary to account for their intensity changes during and following a geomagnetic storm [Farley and Rosen, 1961]. An adequate accelerating mechanism is not known.

In two recent papers [Parker, 1960; 1961] it was shown that violation of neither the third

invariant nor the second or longitudinal invariant can by itself account for the necessary electron acceleration, although the violations contribute to the spatial distribution and decay of all energetic electrons. It was concluded that the necessary electron acceleration must take place in association with violation of the first or transverse invariant $w_{\perp}^2/B = \text{constant}$. The recent work of *Dragt* [1961] has shown that hydromagnetic-wave violation of the first or magnetic moment invariant of energetic protons trapped in the geomagnetic field will produce scattering and loss by collision in the terrestrial atmosphere without changing the proton energy. The same is true for similar first invariant violation of electrons by wave frequencies above the hydromagnetic range. Thus we must inquire into other mechanisms that accelerate electrons at the same time that the invariant is violated.

Therefore in pursuit of further possibilities for the acceleration of Van Allen electrons, we have undertaken to investigate transresonance accelerating. A whistler wave of modest field strength is not sufficient to give the synchronous acceleration of Helliwell and Bell. The wave is too weak to lock in the electrons and carry them to relativistic energies. It is readily shown, however, that synchronism is not necessary for a transfer of energy to the particle. When the wave frequency, as seen in the frame of reference moving with the electron, passes across the electron gyrofrequency, there is a momentary rapid nonadiabatic acceleration of the electron by the wave. The relative phase difference between the electron and the whistler varies in such a way that the acceleration is not reversed, and thereby lost, with the passage of the whistler, as would be the case if resonance were not crossed. The result is a random walk in the particle energy.

To illustrate this transresonant nonsynchronous acceleration, consider the simplest possible case in which a very long wave packet of small and slowly varying amplitude and frequency propagates along a uniform static magnetic field in which there is a charged particle of mass M and charge q . Denote by \mathbf{v} the particle velocity perpendicular to the uniform magnetic field \mathbf{B} ($B \equiv |\mathbf{B}|$). (Motion parallel to the field \mathbf{B} is eliminated by transforming to a coordinate system moving along \mathbf{B} with suitable velocity.) Let the xyz coordinate axes be oriented such

that the z axis of the coordinate system is parallel to and in the same direction as \mathbf{B} . Suppose that there is a plane polarized electromagnetic wave of small amplitude propagating in the z direction with its electric vector \mathbf{e} in the x direction and its magnetic vector \mathbf{b} ($b \equiv |\mathbf{b}|$) in the y direction. Then neglecting terms second order in b/B , the xy motion of the particle is

$$\frac{d^2x}{dt^2} = +\frac{q}{M} E(t) + \Omega \frac{dy}{dt} + O^2 \left(\frac{b}{B} \right)$$

$$\frac{d^2y}{dt^2} = -\Omega \frac{dx}{dt} + O^2 \left(\frac{b}{B} \right)$$

The motion in the z direction is first order in b/B , but it is oscillatory and of no interest in the present problem.

Integrating equation 2, placing the origin of the coordinates so that dy/dt vanishes at $x = 0$ and using the result to eliminate dy/dt in equation 1 yields

$$\frac{d^2x}{dt^2} + \Omega^2 x = \frac{q}{M} E(t)$$

where $E(t)$ is the electric field of the wave at the position of the particle. A discussion of the solution of equation 3 will illustrate the transresonance acceleration.

Let the electric field of the wave packet be confined to the time interval (t_1, t_2) during which the particle is in the wave.

$$E(t) = f(t) \cos \int_{t_1}^t d\tau \omega(\tau)$$

The amplitude $f(t)$ is supposed positive, increasing slowly and smoothly from zero at $t = t_1$ and returning to zero by the time $t = t_2$. The frequency $\omega(t)$ is taken to be a slowly varying function of time throughout (t_1, t_2) and is of the same general order of magnitude as Ω .

The particle motion due to $E(t)$ is

$$\frac{dx_w}{dt} = \frac{q}{M} \int_{t_1}^t d\tau f(\tau) \cos \Omega(t - \tau) \cdot \cos \int_{t_1}^{\tau} du \omega(u)$$

which must be added to whatever motion

$$\frac{dx_i}{dt} = C \cos(\Omega t - \phi)$$

present when $t = t_0$. Now if $f(t)$ and $\omega(t)$ change sufficiently slowly that the amplitude of f changes by only a small fraction in one period, then the solution of equation 5 is the following expression

$$\cong + \frac{qf(t)\omega(t)}{M(\omega^2(t) - \Omega^2)} \sin \int_{t_1}^t d\tau \omega(\tau) \quad (7)$$

Amplitude of the particle motion produced by the wave amplitude $f(t)$ is proportional to $\omega/(\omega^2 - \Omega^2)$ and may be very large if ω is close to Ω . The motion dies away, of course, with the passage of the packet and the eventual return of $f(t)$ to zero. Note however that equation 7 is not valid up to the point that ω becomes equal to Ω , for if $\omega(t) \rightarrow \Omega$ in a finite length of time, then the condition that the rate of change of amplitude be slow is violated. To see what happens in this case, suppose that ω drifts slowly across Ω at some time t_0 during the passage of the wave packet. For simplicity, let t_0 be the origin of the time coordinate so that $t_0 = 0$, and suppose that $\omega(t)$ is changing slowly and linearly with time in the neighborhood of Ω . We write $\omega(t) = \Omega (1 \pm 2\nu^2 \Omega t)$ where ν^2 is small. We are interested in $\omega(t)$ while it is close to Ω , but we suppose that $d\omega/dt$ is so small that ω is close to Ω for very many wave periods. Effectively we are interested in the time interval $|t| \ll 1/\nu^2 \Omega^2$ where $|\Omega t| \gg \nu^2 \Omega^2 t^2 \gg 1$. We assume that $d\omega/dt$ is so small that the change in f may be neglected over the same period. The integral beyond the interval $\pm t_0$ oscillates rapidly that it contributes little to the integral. Thus for convenience we may put the limits equal to $-\infty$, obtaining the integrated motion

$$\cong \frac{qf(0)}{M} \int_{-\infty}^t d\tau \cos \Omega(t - \tau) \cdot \cos (\Omega \tau \pm \nu^2 \Omega^2 \tau^2) \quad (8)$$

If t has become sufficiently large that $\nu^2 \Omega^2 t^2 \gg 1$, the integration is best divided into two intervals, $-\infty$ to $+ \infty$, minus t to 0 . The latter is negligible, yielding

$$\frac{dx_w}{dt} \cong \frac{\pi^{1/2} c}{2\nu} \frac{f(0)}{B} \cos (\Omega t \pm \pi/4) \quad (9)$$

For $\nu \ll 1$, the residual velocity amplitude of transresonance acceleration is large compared to the equivalent velocity amplitude

$cf(0)/B$ of the wave. This velocity may add or subtract from the initial energy, given by equation 6, depending upon the relative phase of dx_i/dt and dx_w/dt . A random walk in energy is the result.

More general calculations show that the same result is reached if both ω and Ω are varying at the position of the particle. The net acceleration is then inversely proportional to the square root of $d(\omega - \Omega)/dt$ instead of just $d\omega/dt$. Extremely large acceleration results in those cases for which the electron motion in the geomagnetic field causes ω and Ω to vary in such a way that $d(\omega - \Omega)/dt$ is very small. The important qualitative distinction to make between transresonant acceleration in the presence of weak whistlers and synchronous acceleration in the presence of strong whistlers is that the latter, with predominantly falling wave frequencies, gives a systematic energy increase, whereas the former gives a random walk. Note that if in some instance rising frequency should be the general rule, then the synchronous acceleration would give a systematic energy decrease, and transresonant acceleration still a random walk.

Conclusion. We suggest transresonant acceleration as a possible contributing factor in the generation of high-energy electrons in the geomagnetic field. Transresonant acceleration affects electrons with initial energies of the order of 10^2 ev or more, which have enough velocity to Doppler-shift the whistler wave frequency in their own frame of reference to the electron cyclotron frequency Ω . The electron velocity change per transresonant passage is of the order of

$$\Delta v \cong c \frac{E}{B} \left| \frac{\Omega^2}{d(\omega - \Omega)/dt} \right|^{1/2}$$

and will be positive or negative according to the relative phase of the whistler and the initial electron motion. On the basis of the simple model presented, the electron velocity will random walk away from the initial velocity by an amount $n^{1/2} \Delta v$ after n transresonant passages, there being no tendency for a systematic velocity increase. The average particle energy will not increase, but the spread in energy will grow with each passing whistler. (This type of random acceleration, in which there may be no change in the mean particle energy, was first discussed by Davis [1956] in connection with Fermi

acceleration of galactic cosmic-ray particles in interstellar space.) A whistler amplitude E of the order of 10^{-3} volts/m (3×10^{-8} statvolts/cm) with $B \sim 10^{-2}$ gauss gives $\Delta v \sim 200$ km/sec, which is about 400 volts for a 500-kev electron. In this case about 10^4 transresonant accelerations would be required to give an rms velocity change of 2×10^9 cm/sec, which could increase 1 kev to 4 kev, or 30 kev to 40 kev.

The relative importance of trans-resonant acceleration depends upon the electromagnetic spectrum in the whistler range in the outer geomagnetic field. The amplitude and the rate of change of frequency, $d(1/\omega)/dt$, must be determined from observation before the net acceleration can be estimated. Whistler fields stronger than 10^{-2} volts/m in combination with a frequency decrease spread over 1 second (or 10^{-1} volts/m and 0.1 second) favor the rapid synchronous acceleration of Helliwell and Bell for electrons that already have near relativistic energies. Nonrelativistic particles and much weaker whistler fields, $E \lesssim 10^{-3}$ volts/m, favor the simple random transresonant acceleration presented here. Intermediate fields enhance the transresonant acceleration for decreasing whistler frequencies by decreasing ν when the phase is such as to increase the particle energy; synchronous acceleration is not achieved, but a velocity increase becomes a little more probable than a decrease, and the number n of trans-resonant passages necessary to achieve a given

velocity increase may be considerably reduced. It will be necessary to carry out a quantitative measurement of whistler amplitudes, etc., in the outer geomagnetic field to assess the effect on the electrons in the outer Van Allen zone.

Acknowledgment. I wish to thank A. J. Dessler for stimulating discussion and many helpful suggestions in the preparation of the manuscript. This work was supported by the National Aeronautics and Space Administration under grant NAS-5-10660, NSG-96-60.

REFERENCES

- Davis, L., Modified Fermi mechanism for acceleration of cosmic rays, *Phys. Rev.*, **135**, 1956.
- Dragt, J. A., Effect of hydromagnetic waves on the lifetime of Van Allen radiation proton, *Geophys. Research*, **66**, 1641-1649, 1961.
- Helliwell, R. A., and T. F. Bell, A new mechanism for accelerating electrons in the outer magnetosphere, *J. Geophys. Research*, **65**, 1839-1848, 1960.
- Parker, E. N., Geomagnetic fluctuations and the form of the outer zone of the Van Allen radiation belt, *J. Geophys. Research*, **65**, 3117-3128, 1960.
- Parker, E. N., Effect of hydromagnetic waves on a dipole field on the longitudinal invariance, *Geophys. Research*, **66**, 693-708, 1961.
- Rosen, A., and T. Farley, Characteristics of the Van Allen radiation zones as measured by a scintillation counter on Explorer VI, *J. Geophys. Research*, **66**, 2013-2028, 1961.

(Manuscript received June 16, 1961.)

A Note on Whistlers Observed above the Ionosphere

J. C. CAIN, I. R. SHAPIRO, J. D. STOLARIK, J. P. HEPPNER

*Space Sciences Division, Goddard Space Flight Center, NASA
Greenbelt, Maryland*

Abstract. The sensing coil of the proton precession magnetometer carried on the Vanguard III satellite also served as an antenna for detecting audio-frequency electromagnetic waves. A preliminary analysis is given for about one hundred whistlers observed September 18 through December 12, 1959, at low latitudes over the altitude range 510 to 3750 km. About 90 per cent of the whistler occurrences are between 6 PM and 6 AM local time indicating a low nighttime absorption by the ionosphere. The intensity of the H component of the wave is estimated to be between 0.01 and 0.5 gamma, with that of a few signals exceeding 1 gamma.

Although the measurements obtained from the Vanguard III satellite (1959 Eta) were primarily in the geomagnetic field [Heppner, Stolarik, Shapiro, and Cain, 1960; Heppner, Cain, Shapiro, and Stolarik, 1961], the sensing coil of the proton precessional magnetometer also served as a magnetic antenna for detecting audio-frequency electromagnetic waves. Detection was possible during each 2-second period following an interrogation on command from the NASA tracking stations. During this interval the coil circuit was connected to an audio-band, high-gain amplifier for transmission of the proton precession signal. There were approximately 100 such transmissions during the 85 days of the satellite life, September 18 to December 12, 1959. The tracking stations were so located that magnetic data were obtained mainly over northern South America, the West Indies, including part of southeastern United States, a small part of southern California, and central Australia. Readings were also taken over southern Australia. Within these regions the orbital inclination of 33.5° gives a geomagnetic latitude range of 15° to 45°. Perigee and apogee altitudes were 510 and 3750 km.

The results presented here are from a preliminary examination of about 100 magnetometer transmissions during which one or more whistlers were recorded. It is estimated that whistlers will be detected on 5 to 10 per cent of the records if a search of all data is complete.

Figure 1 gives typical examples of spectrograms obtained from the magnetometer transmissions. The heavy, horizontal lines are the proton precession signals, $f(\text{cps}) = 4257.6 H$ (gauss),

and their harmonics. The satellite location is given in geomagnetic coordinates, kilometers altitude, and universal time. The value D shown is the whistler dispersion from the well-known approximation [Storey, 1953], $D = dt/df^{-1/2}$, where f is in cycles per second. Figure 1a is an example of a multiple-flash type of whistler group with the small dispersions indicating a single transmission through the primary layers of the ionosphere. Multiple-flash groups are prevalent near the equator with one record showing as many as 20 distinct whistlers within the 2-second interval. The two whistlers (not harmonics) seen in Figure 1b have dispersions of 26 and 24 $\text{sec}^{1/2}$, respectively, and could have come either from two sources or from one source by two distinct paths. Figure 1c illustrates a high dispersion whistler on a record in which the proton precession signal is not visible, probably due to coil orientation relative to the magnetic field.

Figure 2 illustrates the various locations of the satellite for some of the whistlers examined. The numbers given by some of the points indicate their dispersion or, in instances where there are clusters of points, a dispersion representative of the group of whistlers. P and A refer to the heights of perigee and apogee. The geomagnetic longitude range for all the points shown is from 10°E to 15°W except for the two on the far right, which were over Australia. For several of the observations taken at times differing by as much as a month, it has been found that the ratio of the measured dispersion to the path length along a line of force to the ground was constant. Sample calculations have been made

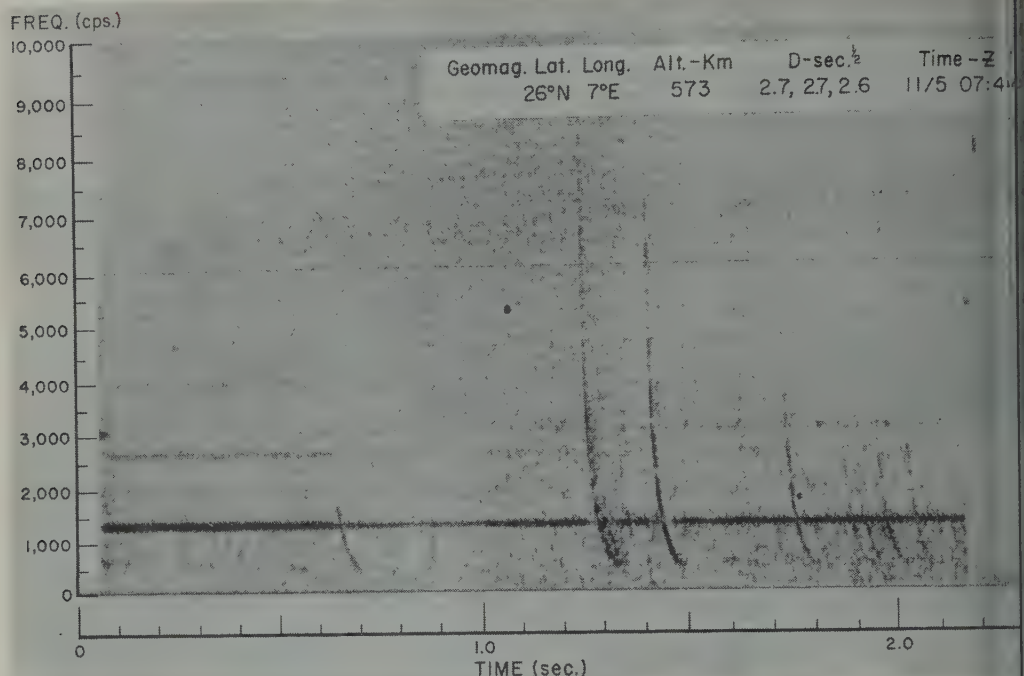


Fig. 1a. Whistler spectrogram.

using average electron densities for a quiet, nighttime ionosphere [Wright, Wescott, and Brown, 1959] and a dipole field extrapolation in the simplified formula [Helliwell and Morgan, 1959]:

$$D = \frac{1}{2c} \int_{\text{path}} \frac{f_p}{\sqrt{f_H}} ds$$

For a wave traveling from the surface to 550 km

along the field line originating at 35°N geomagnetic latitude, the calculated result of $D = 2.7$ sec^{1/2} is within about thirty per cent of values shown in Figure 2 near that location. Of those points near perigee from about 1 to 35°N, it is noted that the two for which $D = 2.7$ and 5.0 were observed during magnetically quiet periods, whereas the rest ($D \leq 2.8$) were obtained at times of magnetic disturbances.

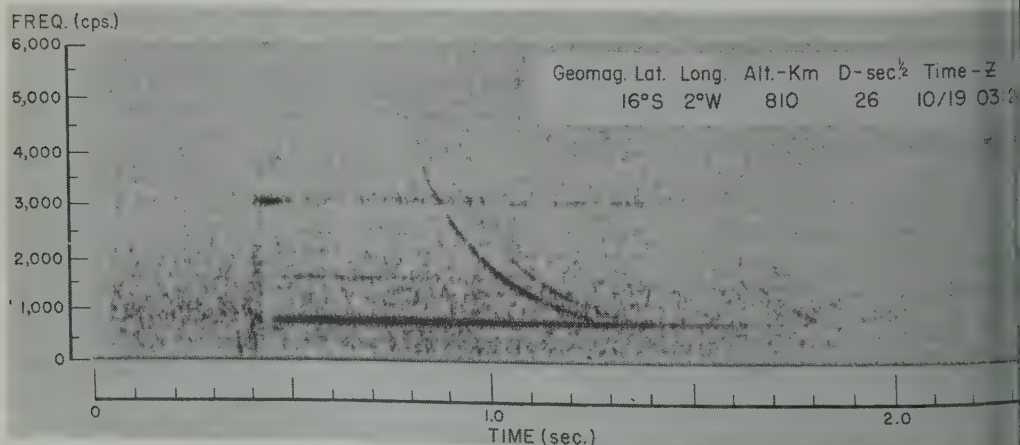


Fig. 1b. Whistler spectrogram.

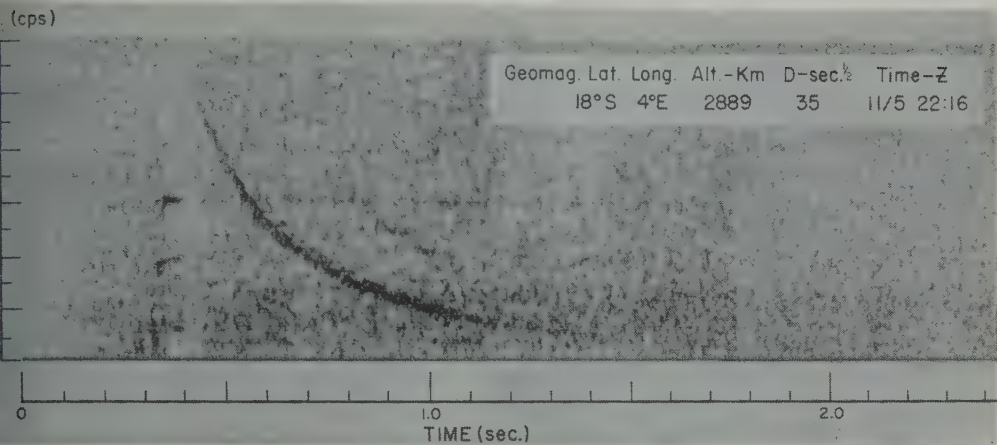


Fig. 1c. Whistler spectrogram.

though the survey of all the Vanguard III for whistlers is yet incomplete, some statements can be made concerning the statistics of reference on the assumption that those investigated constitute a random sample. It is found that 60 per cent of the whistlers examined during the 25 days listed as 'storm' by three or more magnetic observatories (Mann, 1960; 1961). A strong diurnal variation is noted with about 90 per cent of the

whistlers occurring during the interval from 6 PM to 6 AM local time.

Precise data are not available on signal amplitudes as no in-flight calibrations were provided for the amplitude response of the instrument. Measurements of the relative amplitudes between the whistler and proton precession signal on a particular record cannot be used to compute the absolute whistler amplitude owing to the unknown angles between the coil axis,

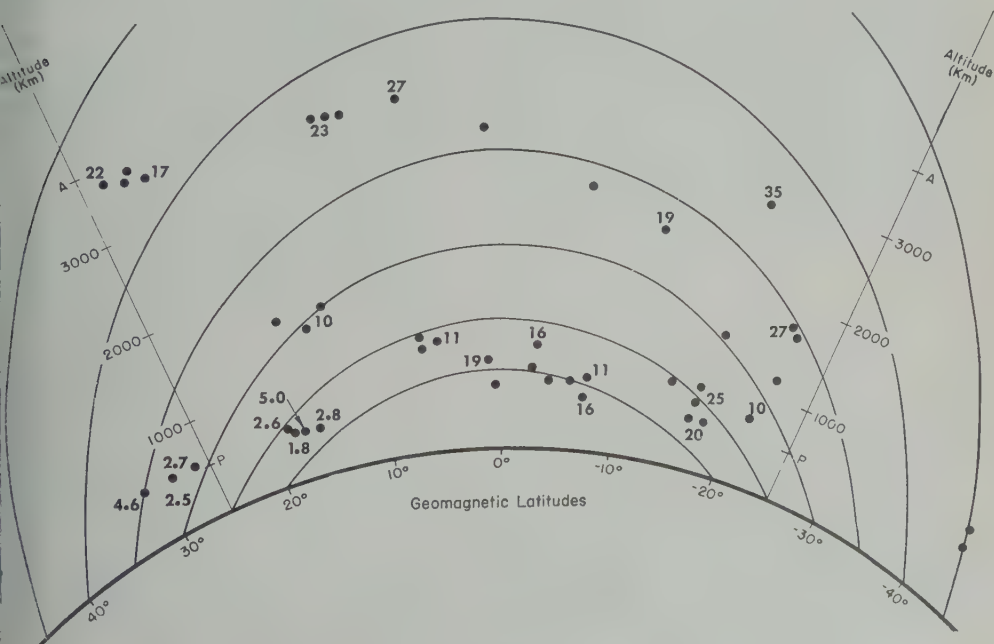


Fig. 2. Whistler dispersions (sec^\dagger).

magnetic field, and the whistler wave normal. However, one can make approximate estimates based on the maximum proton signal strengths and the detection limitations of the instrument. The 600-turn, 6-millihenry coil and amplifier with a voltage gain of about 10^7 had a maximum response at a frequency of 1 kc/s. At 0.35 and 4 kc/s the sensitivity was down by a factor of 10, and at 10 kc/s the sensitivity was down by a factor of more than 100 relative to that at 1 kc/s. A computation reveals that nearly all of the observed whistlers correspond to VLF waves whose H component lies between about 0.001 and 5 gammas ($1 \text{ gamma} = 10^{-9} \text{ weber/m}^2$). It is thought that most of those examined have an amplitude between 0.01 and 0.5 gamma with a few signals exceeding 1 gamma. The fact that frequencies as high as 10 kc/s are observed supports this belief.

Pending further analysis a few important inferences may be drawn from the data presented here. We have adopted the hypothesis that these observed signals are in fact 'fractional-hop' whistlers propagating from lightning strokes near the earth's surface to the satellite along geomagnetic lines of force. The statistics of occurrence imply that the ionization becomes more field aligned during magnetic disturbance. Further work involving computations of dispersions using the best available electron densities at the time of each observation is being done to demonstrate the hypothesis in detail. If the assumption seems justified, the analysis will provide direct information on the electron densities above the F region.

Both the strength of the received signals and the diurnal variation of whistler occurrence

imply that the nighttime attenuation of the ionosphere is not high. This conclusion is confirmed by the 18 kc/s measurements made with the Lofti satellite [Beece and Young, 1961]. It thus appears that the high daytime absorption and nighttime transparency of the ionosphere are features common to the frequency band from 18 kc/s below 45° geomagnetic latitude. The signal strength and the relatively high occurrence near the equator suggest that whistlers may be of importance in the acceleration and diffusion of particles in the inner radiation belt [Adams, 1959; Herlofson, 1960; Helliwell and Bell, 1961].

REFERENCES

- Alfvén, H., *Phys. Rev. Letters*, **3**, 459-460, 1961.
 Beece, L. S., and C. E. Young, Paper presented at the URSI-IRE joint Meeting, Washington, D. C., May 1961.
 Helliwell, R. A., and T. F. Bell, *J. Geophys. Research*, **65**, 1839, 1960.
 Helliwell, R. A., and M. G. Morgan, *Proc. IRE*, **200**-208, 1959.
 Heppner, J. P., J. D. Stolarik, I. R. Shapiro, and J. C. Cain, *Space Research*, First International Space Science Symposium, edited by H. K. Mann Bijl, North-Holland Publishing Company, Amsterdam, 1960.
 Heppner, J. P., J. C. Cain, I. R. Shapiro, and J. D. Stolarik, *NASA Tech. Note D-696*, 1961.
 Herlofson, N., *Phys. Rev. Letters*, **5**, 414-416, 1960.
 Lincoln, J. Virginia, *J. Geophys. Research*, **65**, 797, 1323-1326, 1960.
 Lincoln, J. Virginia, *J. Geophys. Research*, **66**, 662, 1282, 1961.
 Storey, L. R. O., *Phil. Trans. Roy. Soc. A*, **253**, 113-141, 1953.
 Wright, J. W., L. R. Wescott, and D. J. Brown, *NASA Tech. Note No. 40-3*, 1959.

(Manuscript received June 9, 1961.)

Drift Rate in a Dipole Field

JOHN S. LEW

*Physics Division, Research Directorate, Air Force Special Weapons Center
Air Research and Development Command, Kirtland Air Force Base, New Mexico*

Abstract. Expressions are found for the drift of charged particles in the field of a magnetic dipole. These give, in the guiding center approximation, the longitude drift rate and drift period of relativistic particles with mirror points at any latitude. The results are applied to trapped electrons and protons in the earth's magnetic field, and are adapted for further use in such applications.

Introduction. We calculate here in greater detail than in previous papers [Welch and Whittaker, 1959; Hamlin, Karplus, Vik, and Watson, 1960] the drift rate of a charged particle in the field of a magnetic dipole. To do this, we take the drift velocity of the guiding center in its relativistic form, find the resulting drift in longitude and average over one bounce period. For a particle and a fixed dipole, this rate is the sum of a constant, and terms depending on the energy, the equatorial radius, and the mirror-point latitude, respectively, of the particle in its orbit. We apply this expression to protons and electrons in the earth's field, and calculate the drift period for each particle, that is, the time required for each to drift once around the earth. Our conclusions are summarized in two graphs and agree with the earlier work of Welch and Whittaker.

1. Motion in a general irrotational field. Consider a particle of charge q and rest mass m , moving in a magnetic field \mathbf{B} , which we assume to be that of a dipole, and in an electric field \mathbf{E} which we take throughout to be zero. We let

\mathbf{v} = velocity of particle

$\beta = v/c$

$\gamma = (1 - \beta^2)^{-1/2}$

\mathbf{e} = unit vector by \mathbf{e} and s , respectively, a unit length coordinate along the field line, and resolve \mathbf{v} parallel and perpendicular

to \mathbf{B} . If $\mathbf{E} = 0$, the relativistic motion of our particle is described by

$$\left(\frac{d}{dt}\right)(m\gamma\mathbf{v}) = \frac{q}{c}\mathbf{v} \times \mathbf{B} \quad (1)$$

from which it follows that v^2 is conserved, or that v , β , and γ are constant. Hence the motion is just that of a Newtonian particle with effective mass $m\gamma$. In the guiding center approximation, which we assume valid, the guiding center of this particle moves generally back and forth along a field line, while drifting slowly to one side. In fact, according to the previous remark, we can take the 'Newtonian' equations, derived rigorously by Bogolyubov and Zubarev [1955], for the motion of the guiding center, and simply replace m by $m\gamma$. The drift velocity is therefore given by

$$\mathbf{v}_d = \frac{m\gamma c}{qB} \mathbf{e} \times \left(\frac{1}{2} v_{\perp}^2 \frac{\text{grad } B}{B} + v_{\parallel}^2 \frac{\partial \mathbf{e}}{\partial s} \right) \quad (2)$$

Now $\partial \mathbf{e} / \partial s = \kappa \mathbf{n}$, where κ and \mathbf{n} are the curvature and unit normal, respectively, of the field lines. Moreover, if $\text{curl } \mathbf{B} = 0$, which is true in particular of a dipole field, then

$$\frac{\text{grad } B}{B} = \frac{\mathbf{e}}{B} \frac{\partial B}{\partial s} + \kappa \mathbf{n}$$

Substituting in (2) we see that

$$\mathbf{v}_d = (m\gamma c \kappa / qB) \left(\frac{1}{2} v_{\perp}^2 + v_{\parallel}^2 \right) \mathbf{b} \quad (3)$$

where $\mathbf{b} = \mathbf{e} \times \mathbf{n}$ is the binormal to the magnetic field lines. But, by an earlier statement,

$$v^2 = v_{\parallel}^2 + v_{\perp}^2 = v_{\parallel}^2 + 2B\mu = \text{constant} \quad (4)$$

at every point of the particle's orbit, where $\mu = v_{\perp}^2 / 2B$ is a constant in this approximation. If we evaluate μ at the mirror points, at which $v_{\parallel} = 0$, and substitute in (4), then we find

$$v_{\parallel}^2 = v^2 (1 - (B/B_m)) \quad (5)$$

and, putting (4) and (5) into (3), we get at last

$$\mathbf{v}_d = \left(\frac{mc^3}{2q} \right) \beta^2 \gamma \left(\frac{\kappa}{B} \right) \left(2 - \frac{B}{B_m} \right) \mathbf{b} \quad (6)$$

Longitude drift in a dipole field. Now suppose the field \mathbf{B} is produced by a dipole of moment \mathbf{a} . We introduce a system (r, θ, ϕ) of spherical coordinates, with the z axis along \mathbf{a} , and define the latitude λ as $\frac{1}{2}\pi - \theta$. In these coordinates, each field line has an equation

$$r = r_0 \cos^2 \lambda \quad (7)$$

where r_0 is a constant, the value of r at $\lambda = 0$. The curvature κ and magnetic field B at a point on this field line are then given by

$$\kappa = \frac{3}{r_0 \cos \lambda} \frac{1 + \sin^2 \lambda}{(1 + 3 \sin^2 \lambda)^{3/2}} \quad (8)$$

$$B = (a/r_0^3) \cos^{-6} \lambda (1 + 3 \sin^2 \lambda)^{1/2} \quad (9)$$

We could put these equations in (6) to get \mathbf{v}_d ; but the quantity of greater interest is $\dot{\phi}$, the rate of change of longitude. Since the binormal \mathbf{b} is perpendicular to the plane of the field line, the entire magnitude of \mathbf{v}_d goes into changing ϕ . Thus

$$\dot{\phi} = v_d / r \cos \lambda = v_d / r_0 \cos^3 \lambda \quad (10)$$

When we apply this formula to the earth, it is useful to let r_e be the earth's radius. From equation 9, $B_e = a/r_e^3$ is then the field at $r = r_e$, $\lambda = 0$, that is, the field at the earth's surface on the magnetic equator. Thus (6)–(10) together yield

$$\dot{\phi} = K \beta^2 \gamma \left(\frac{r_0}{r_e} \right) \frac{(1 + \sin^2 \lambda) \cos^2 \lambda}{(1 + 3 \sin^2 \lambda)^2} \cdot \left(2 - \frac{B}{B_m} \right) \quad (11)$$

where we have defined

$$K = 3mc^3 / 2qB_e r_e^2 \quad (12)$$

For an arbitrary dipole, we can still arrive at equations 11 and 12 if we let r_e be some other standard length. Then for each given particle, with a fixed value of q/m , K is a constant with dimensions (time)⁻¹.

A proton or electron makes many bounces in drifting around the earth. Hence we can replace $\dot{\phi}$, the instantaneous drift, by $\langle \dot{\phi} \rangle_{av}$, the average drift in a bounce period P . Now $v_{||}$ is the velocity of the guiding center, as well as the particle,

along a field line; so

$$\langle \dot{\phi} \rangle_{av} = \frac{1}{P} \int_0^P \dot{\phi} dt = \frac{1}{P} \oint \frac{\dot{\phi}}{v_{||}} \frac{ds}{d\lambda} \cdot d\lambda$$

$$P = \int_0^P dt = \oint \frac{1}{v_{||}} \frac{ds}{d\lambda} d\lambda$$

Using (5) and (7) to express $v_{||}$ and we find from (11)–(14)

$$\langle \dot{\phi} \rangle_{av} = K \beta^2 \gamma \left(\frac{r_0}{r_e} \right) \frac{F(\lambda_m)}{G(\lambda_m)}$$

where λ_m is the mirror-point latitude, and G , with $x = \sin \lambda$, are given by

$$F(\lambda_m) = \oint \frac{1 - x^4}{(1 + 3x^2)^{3/2}} \cdot \frac{2 - B/B_m}{(1 - B/B_m)^{1/2}} dx$$

$$G(\lambda_m) = \oint \frac{(1 + 3x^2)^{1/2}}{(1 - B/B_m)^{1/2}} dx$$

For large γ , i.e. relativistic energies, factor $\beta^2 \gamma$ approaches γ ; in fact, $\beta^2 \gamma \rightarrow (1 - \gamma^{-2})$. For small λ_m , i.e. mirror point the equator, the integrands of equations 15 and 17 are nearly equal; in fact, all terms $(1 - B/B_m)^{1/2}$ approach unity. Thus for large γ and small λ_m we can approximate

$$\langle \dot{\phi} \rangle_{av} \sim K \gamma (r_0 / r_e)$$

Values of $F(\lambda_m)$ and $G(\lambda_m)$. The values of F and G are pure numbers, independent of everything but the mirror-point latitude. We shall discuss these functions before we discuss the special problem of drift around the earth. We shall calculate F and G at the extreme limits $\lambda_m = 0$ and $\lambda_m = \pi/2$, and then give the results obtained by computer for intermediate λ_m .

First we refine the small λ_m approximation used in (18). Setting $B_0 = a/r_0^3$ we expand

$$B = B_0 \left(1 + \frac{9}{2} x^2 + \frac{75}{8} x^4 + \frac{277}{16} x^6 + \dots \right)$$

To simplify the factor $(1 - B/B_m)^{-1/2}$ in (15) and (17), we then define a new variable y by

$$B = B_0 \left(1 + \frac{9}{2} y^2 \right), \quad y/x > 0$$

TABLE 1. Values of F , G , F/G , and G/F as a function of mirror-point latitude λ_m . The quotient F/G (drift rate of particles mirroring at λ_m)/(drift rate of particles mirroring on the equator). The quotient G/F = (drift period of particles mirroring at λ_m)/(drift period of particles mirroring on the equator).

F	G	F/G	G/F
2.96162	2.96162	1.00000	1.00000
2.98273	3.00483	.99265	1.00741
3.03940	3.12555	.97243	1.02835
3.12360	3.31013	.94365	1.05972
3.22309	3.53927	.91067	1.09810
3.32544	3.79349	.87662	1.14075
3.42050	4.05604	.84331	1.18580
3.50152	4.31380	.81170	1.23198
3.56538	4.55732	.78234	1.27822
3.61195	4.78010	.75562	1.32341
3.64322	4.97786	.73188	1.36634
3.66228	5.14765	.71145	1.40558
3.67255	5.28740	.69459	1.43971
3.67719	5.39553	.68152	1.46730
3.67872	5.47081	.67243	1.48715
3.67899	5.51280	.66735	1.49846
3.68054	5.52081	.66667	1.50000

vious but straightforward algebra, we have for x ,

$$\left(1 - \frac{25}{24}y^2 + \frac{2159}{1152}y^4 + \dots\right) \quad (21)$$

write both integrals in terms of y , and take the first few terms. The resulting two in y_m we express in terms of x_m (subscript m as value at mirror point), and find that

$$\frac{2\pi\sqrt{2}}{3} + \left(\frac{11}{16}x_m^2 - \frac{791}{1024}x_m^4 + \dots\right) \quad (22)$$

$$\frac{2\pi\sqrt{2}}{3} - \left(\frac{23}{16}x_m^2 - \frac{1231}{1024}x_m^4 + \dots\right) \quad (23)$$

for any γ and small λ_m we conclude that

$$K\left(\gamma - \frac{1}{\gamma}\right)\left(\frac{r_0}{r_e}\right) - \frac{3}{4}x_m^2 + \frac{193}{128}x_m^4 + \dots \quad (24)$$

Next we estimate F and G when λ_m is near $\frac{1}{2}\pi$. Now B contains a factor $\cos^{-6}\lambda$, and thus grows rapidly as λ increases. Hence, for $\lambda_m \sim \frac{1}{2}\pi$, the quantity B/B_m is small over most of the integration, and we may approximate B/B_m by the constant B_0/B_m . Then (16) and (17) become standard integrals, which we can evaluate and then expand in $w_m = \cos \lambda_m$; this yields

$$F(\lambda_m) \sim 2\alpha + O(w_m^4) \quad (25)$$

$$G(\lambda_m) \sim 3\alpha - 4w_m^2 \quad (26)$$

where

$$\alpha = (2/3\sqrt{3})(2\sqrt{3} + \log_e(2 + \sqrt{3})) = 1.84027$$

Thus for any γ and $\lambda_m \sim \frac{1}{2}\pi$ we conclude that

$$\langle\phi\rangle_{av} \sim K\left(\gamma - \frac{1}{\gamma}\right)\left(\frac{r_0}{r_e}\right)\left(\frac{2}{3} + \frac{8}{9\alpha}w_m^2\right) \quad (27)$$

When $\lambda_m = \frac{1}{2}\pi$ precisely, then B_m is infinite, and the approximation becomes exact. That is, a particle with mirror points at the poles drifts $2/3$ as fast as a particle of the same γ and r_0 , with mirror points at the equator.

The intermediate values of F and G were numerically integrated on a computer by Duane C. Jensen, and used to calculate F/G and G/F . The resulting numbers are given in Table 1 and

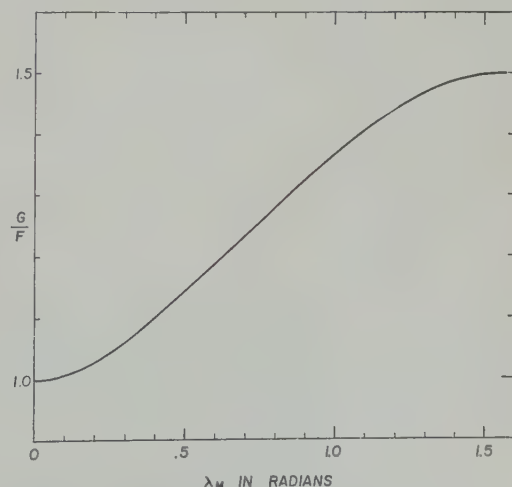


Fig. 1. Latitude quotient G/F as a function of mirror-point latitude λ_m . The quotient G/F = (drift period of particles mirroring at λ_m)/(drift period of particles mirroring on the equator).

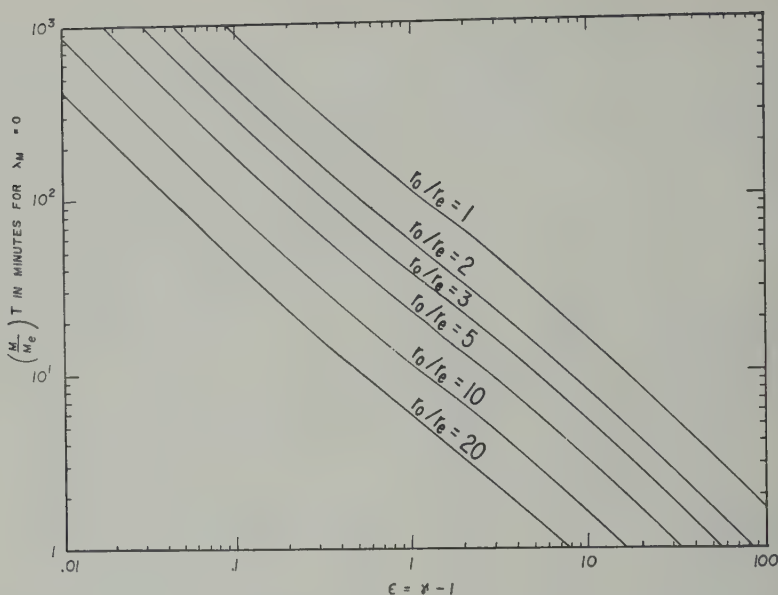


Fig. 2. (Drift period of particles mirroring on the equator) \times (mass in electron masses) as a function of $\epsilon = \gamma - 1 =$ particle kinetic energy/rest energy, for several values of $r_0/r_e =$ field-line equatorial radius/earth's radius. See text to find the values used for the earth's radius and surface equatorial field.

are probably accurate to better than 10^{-3} . The quantity G/F (which is the ratio of drift period for mirror latitude λ_m to drift period for mirror latitude 0) has been plotted against λ_m in Figure 1. An expression which fits this graph to better than 10^{-3} is

$$G/F = A + Bx_m^2 - C \exp(-kx_m^2) \quad (28)$$

in which the best values of the constants are

$$A = 1.04675$$

$$B = .45333$$

$$C = .04675$$

$$k = 6.34568$$

These numbers were also obtained by Mr. Jensen.

Drift time in the earth's field. For protons or electrons in the earth's field, we can take

$$q = 4.774 \times 10^{-10} \text{ esu}$$

$$r_e = 6.378 \times 10^8 \text{ cm}$$

$$B_e = .312 \text{ gauss}$$

and therefore

$$qr_e B_e = 9.501 \times 10^{-2} \text{ erg} = 59.30 \text{ b}$$

If we denote the electron mass by m_e , then have

$$K = \left(\frac{m}{m_e}\right) \frac{3}{2} \frac{m_e c^2}{qr_e B_e r_e} c$$

$$= \left(\frac{m}{m_e}\right) \times 6.075 \times 10^{-4} \text{ sec}^{-1}$$

It is convenient to express $\langle \phi \rangle_{av}$ in terms of kinetic energy; so we define

$$\epsilon = \gamma - 1 = \text{kinetic energy}/mc^2$$

Putting these equations in (18), we see that

$$\langle \phi \rangle_{av} = 6.075 \times 10^{-4} \text{ sec}^{-1}$$

$$\frac{\epsilon(2 + \epsilon)}{1 + \epsilon} \frac{m}{m_e} \frac{r_0}{r_e} \frac{F}{G}$$

The most interesting number to compute is the time T required to drift once around the earth. Clearly, $T = 2\pi/\langle \phi \rangle_{av}$; so we find

$$T = \frac{172.4 \text{ minutes}}{\epsilon} \frac{1 + \epsilon}{2 + \epsilon} \frac{m_e r_e}{m r_0} \frac{G}{F}$$

At high and low energies, we can approx

for $(1 + \epsilon)/(2 + \epsilon)$, and further simplify thus within 2 per cent

$$\frac{36.2 \min}{\epsilon} \frac{m_e r_e G}{m r_0 F} \quad \text{for } \epsilon < .04 \quad (33)$$

$$\frac{172.4 \min}{\epsilon} \frac{m_e r_e G}{m r_0 F} \quad \text{for } \epsilon > 50 \quad (34)$$

electron and proton of the same kinetic energy have comparable drift periods, since the $m\epsilon$ in the denominator of (32) is the same for both; the proton is always faster, but more than twice as fast. For example, take a 1 Mev proton and electron at 2 earth radii, mirroring on the equator; then the proton period is 45 sec, and the electron period is 87 sec. At intermediate energies we have plotted the drift period in figure 2; that is, we plotted $(m/m_e) T$ vs. ϵ for $\lambda_m = 0$ and for various choices of (r_0/r_e) . When the drifting particles are electrons, the graph yields T from (kinetic energy in Mev/.511). When the drifting particles are protons, the

graph yields 1836 T from (kinetic energy in Mev/931).

Acknowledgments. The author has received considerable help from Captain Jasper A. Welch, who proposed this discussion and made numerous suggestions to improve it; and from Mr. Duane C. Jensen, who carried out the numerical part of this work on a computer.

REFERENCES

- Bogolyubov, N. N., and D. N. Zubarev, An asymptotic approximation method for a system with rotating phases and its application to the motion of a charged particle in a magnetic field, *Ukrainian Math. J.* 7 (5), 1955, English translation by Burton D. Fried, *Space Technology Laboratories, Inc.*, January 1960.
- Hamlin, D. A., R. Karplus, R. C. Vik, and K. M. Watson, Mirror and azimuthal drift frequencies for geomagnetically trapped particles, *Convair, San Diego Physics Section*, 1960.
- Welch, J. A., Jr., and W. A. Whittaker, Theory of geomagnetically trapped electrons from an artificial source, *J. Geophys. Research*, 64, 909-922, 1959.

(Manuscript received January 23, 1961.)

Rocket Measurements of the Magnetic Field above New Mexico

R. HUTCHINSON AND B. SHUMAN

*Geophysics Research Directorate
Air Force Cambridge Research Laboratories
Bedford, Massachusetts*

Abstract. Absolute value, total intensity measurements of the magnetic field above Holloman Missile Development Center, New Mexico, were obtained to an altitude of 236 km with a proton precession magnetometer aboard an Aerobee-Hi rocket. The flight took place on a magnetically quiet day during the normal midmorning variation. Comparison of the observed values with those predicted by spherical harmonic coefficients yields an anomaly of the order of 80 γ during ascent between the altitudes of 85 and 165 km, whereas no comparable effect was found for the descent portion of the flight down to 146 km.

Introduction. Studies of ionospheric current densities and magnetic fields surrounding the earth, with data from rocket-borne magnetometers, will yield information and problem solutions not possible from surface measurements.

In the first attempts to measure the magnetic field from rockets a Naval Ordnance Laboratory proton magnetometer was used [Maple, Singer, and Singer, 1950; Singer, Maple, and Singer, 1951]. No evidence of magnetic fields was observed by upper-atmosphere current sheets was obtained during an early afternoon firing at Sands Proving Ground, New Mexico. The observed experimental data and an inverse cube law agreed to within 100 γ at altitudes above 100 km. Later firings, conducted from the U.S.S. *Albatross* off the coast of Peru near the magnetic equator, showed a field decrease in accordance with a simple dipole for an early evening firing, and a discontinuity of several hundred gammas in the altitude range 93 to 105 km local noon firing.

Stolarik, and Meredith [1958] compared these studies with the firing at WSPG of an Aerobee containing a proton precession magnetometer, and with subsequent firings from the facility at Fort Churchill, Canada. Marks and Stanley [1959] also utilized this site, firing a Nike-Cajuns equipped with similar magnetometers, and Cahill [1959] completed a series of measurements using balloon-launched rockets fired off the coast of Greenland, Christmas Island in the mid-Pacific, and in the South Sea off Antarctica.

Flight objectives. The firing at Holloman Missile Development Center (HMDC) was carried out as part of the normal rocket research program being conducted by the Ionospheric Physics Laboratory at Geophysics Research Directorate. The purpose of this experiment was to explore the current system of the daily magnetic variations as evidenced by the noon minimum during an interval of magnetic calm. The current system at near midnight was previously investigated by the 1956 WSPG firing. Chapman [1954], in suggesting a group of experiments for rocket magnetic measurements of the ionosphere, called for similar investigations.

A Varian ground station proton precession magnetometer was installed in the Aerobee blockhouse, and the sensing coil was set about 60 meters from the blockhouse and about 160 meters from the launching tower. Records of total intensity versus time were obtained for a period of approximately 2 months prior to the firing; they showed that the diurnal dip at HMDC averaged about 30 γ and reached a maximum between 10 and 12 A.M., MST. Figure 1 shows the daily trend in F for several days before the firing and through the actual flight period. The actual launching time, 0940 MST, June 10, 1959, was a compromise between the time of maximum dip and the time for the most favorable winds aloft over the range.

Rocket installation and instrumentation. The vehicle used was an Aerobee-Hi model AJ 11 with parachute, beacon, and severance ring extension sections. The ogive was Fiberglas, and the instrumentation rack was constructed of

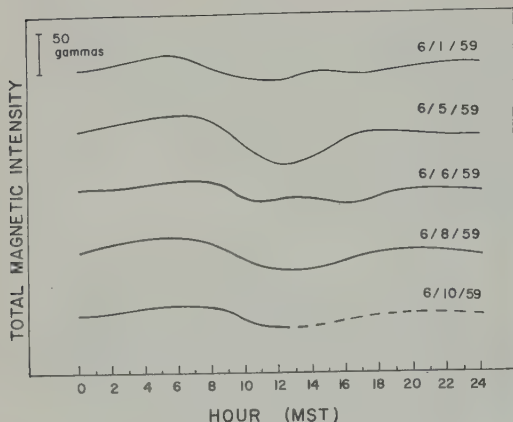


Fig. 1. Total magnetic intensity at HMDC launching site.

wood. A series of ground swings made at the location of the station magnetometer indicated that the total magnetic disturbance of the vehicle in its flight orientation would be $+20 \gamma$ at the position of the rocket magnetometer sensing coils. Of this, about 12γ would be permanent magnetism and 8γ induced magnetism.

Instrumentation consisted of a Varian proton precession magnetometer [Packard and Varian, 1954; Packard, 1954] with dual sensing coils, one mounted with its coil axis horizontal and one vertical; three Schonstedt magnetic aspect sensors mounted orthogonally; a 10-watt Fm-Fm transmitter with fin notch antenna and four subcarrier oscillators; and a Miran beacon assembly. Power for the units was obtained from silver cells and from series-parallel strings of radio B batteries; no d-c to d-c converters were used. All possible precautions were taken to prevent introduction of magnetic material into the nose cone. Nylon bolts and Fiberglas hold-down straps were generally used in mounting the instrumentation, but a few aluminum and brass bolts were used at maximum strain points. The nose cone ogive usually constructed of aluminum, was made of Fiberglas to prevent the polarizing current surges from introducing eddy currents into the shell that would interfere with the desired proton precession decay characteristics. Figure 2 shows the rack assembly and the nose cone.

Tests performed on the nose cone, rack, and extension sections at the Air Force Magnetic Test Station, Weston, Massachusetts, indicated

that the total disturbance field of the assembly was less than 3γ and the variation in this disturbance as the unit was rotated about its longitudinal axis was less than 1γ .

Flight summary. The position of the vehicle as a function of time was determined very

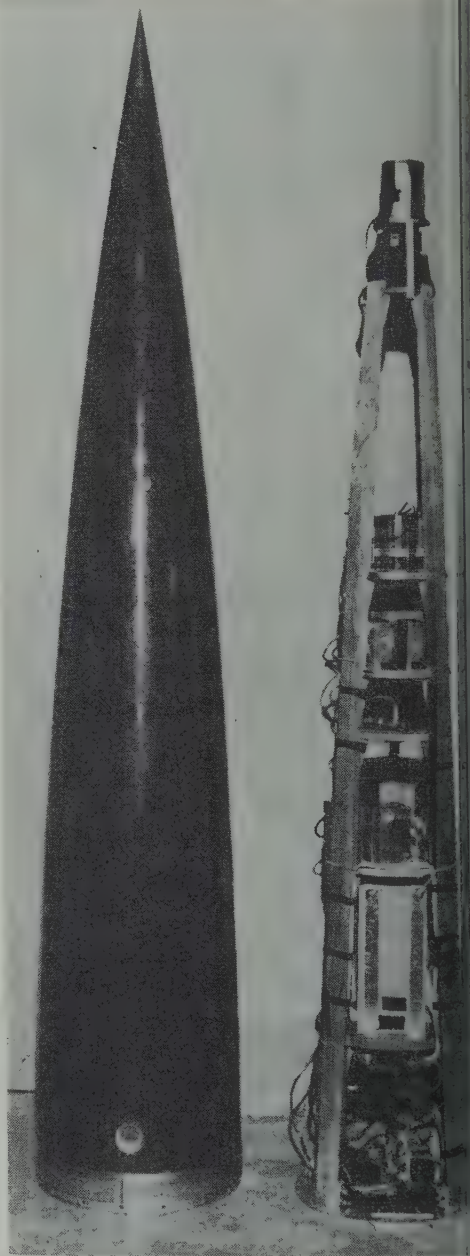


Fig. 2. Nose cone rack assembly and Fiberglas nose cone.

by several independent systems. Throughout the flight, data were obtained from cine-strip cameras, fixed cameras, velocimeters, and the Miran beacon system provided good data until re-entry and breakup, and the radar provided data from take-off to re-entry. Plots of X (X axis is positive true north) and of Y (Y axis is positive east) versus altitude were essentially straight lines from burnout to re-entry. At peak altitude of 236 km, the ground track along the ground was 28 km to the northwest.

Altitude of the vehicle as a function of time was obtained from the aspect sensor data. A constant roll rate of about 2 rps was achieved after take-off. It was also determined that the angle between local vertical and the vehicle's figure axis did not vary more than $\pm 1^\circ$ during the flight.

The magnetometer output was recorded on tape at several receiving sites both as a continuous signal and as the demodulated precession signal together with a reference standard frequency. In addition, the precession frequency was recorded on a high-speed oscillograph which was also furnished to the blockhouse for monitoring. A total of 728 precession signals was recorded at a rate of approximately 2 per second, recorded as a function of time, each signal lasting about $\frac{1}{4}$ second long. During any $\frac{1}{4}$ -second period of flight, the rocket's ground velocity was not greater than 6200 m/sec and the maximum gradient was less than 1 mG. Therefore, the maximum error that could result from averaging the value of precession frequency over a period of flight of $\frac{1}{4}$ seconds is $\pm 6 \gamma$, and the value of this error at re-entry would be less than $\pm 1 \gamma$.

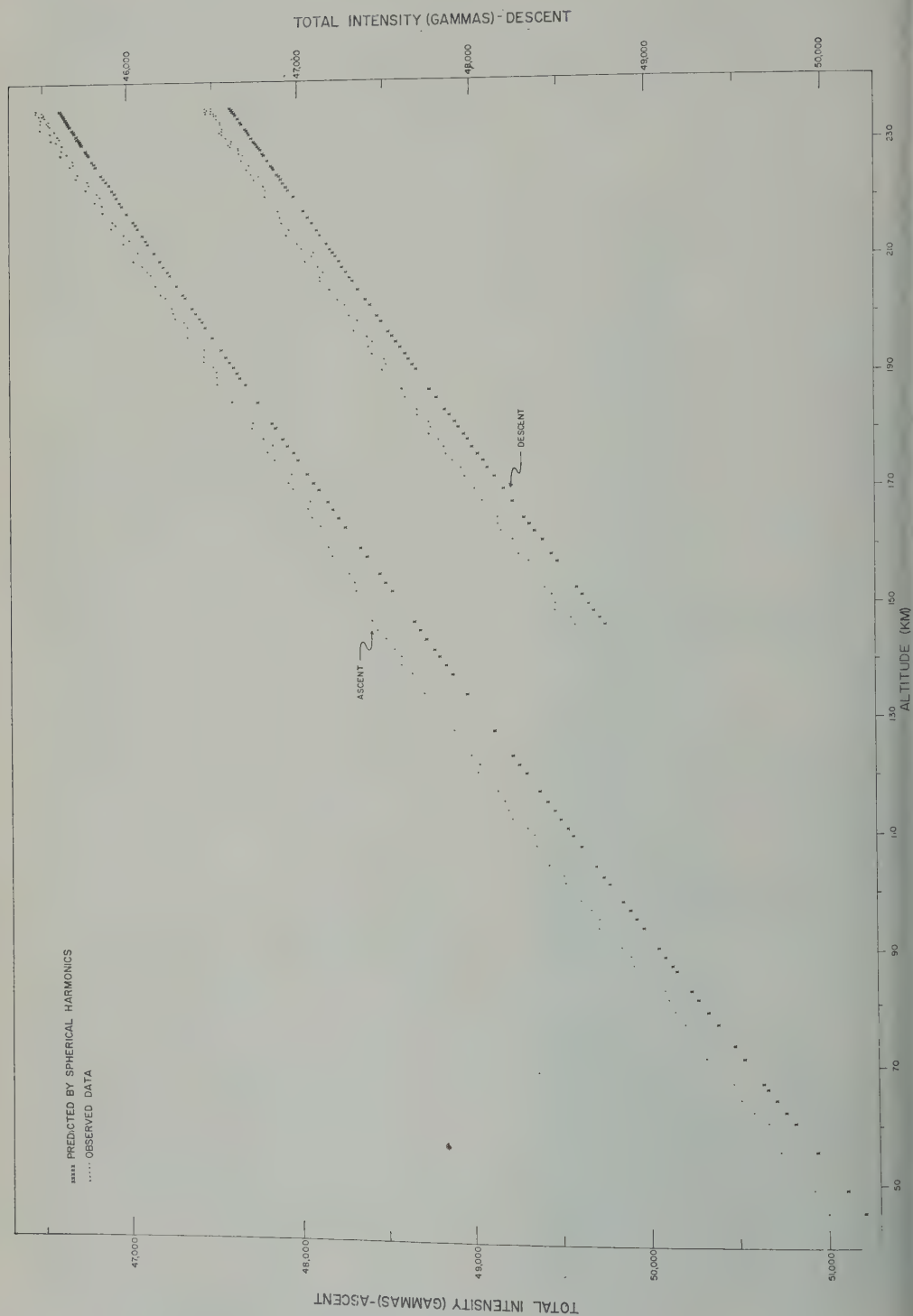
Reduction and analysis. The scalar value of the total magnetic intensity F is related to the proton precession frequency f by the expression $\omega = \gamma_p F$, where $\omega = 2\pi f$, and γ_p , the gyromagnetic ratio of the proton, has a numerical value of $2.67513 \pm .00002 \times 10^4$ cgs. Therefore, $F = 2\pi f / \gamma_p = 23.4874 f \gamma$. Counting this frequency directly over the short time interval of one precession period will not yield f to an accuracy of better than 1 part in 500, so other techniques were utilized. The taped demodulated precession frequency was filtered and then gated by the gating signal in a Dynac 5095 counter combination. The gate signals were

used in turn to control the passage of a number of cycles of a much higher reference frequency into a digital counter section, and by using the 50 kc/s standard taped on the original records as the reference frequency, no special tape speed compensation was required.

If f is the precession frequency in cps, n = number of counts that the gate is preset to pass, N = displayed number of cycles of the reference frequency passed by the gate, and with a reference frequency of 50 kc/s, the gate is open n/f seconds and $f = 5 \times 10^4 n/N$. With $n = 200$ cycles and N being in the range of 4500 to 5000 counts, the ± 1 count uncertainty in the frequency counter now represents an error of $\pm 10 \gamma$.

Of the original 728 readings, only 230 were found to be consistently repeatable to within one count on all four readout runs. Almost all of these 230 values of f were obtained from the horizontal coil. This was not unexpected inasmuch as the output signal of the magnetometer is proportional to $\sin^2 \theta$, where θ is the included angle between the coil axis and the earth's field vector. Also, there were no repeatable values during the powered portion of the rocket flight and it is believed that this may be due to vibrational effects on the system. The frequency values were then converted to total intensity in γ , subject to corrections for the rocket roll rate [Bloom, 1955] and the permanent and induced field of the rocket body. Inasmuch as the rocket roll was a steady 2 rps throughout the data-collecting portion of the flight, this correction was a constant.

For the position of each of the 230 valid data points, an extrapolated total intensity value was calculated by using the spherical harmonic coefficients of Finch and Leaton [1957]. A plot of this result is compared with the (uncorrected) rocket data in Figure 3. Figure 4 shows the difference between the observed and calculated values of the total intensity at each point. This latter result, smoothed by running averages of 5, is also shown in Figure 4. This difference for the ascent curve is a fairly straight line except for the anomalous section between 85 and 165 km. The magnitude of the anomaly is more than 80 γ and is negative in the sense that the observed field is less than was expected. (The experimental error is considered to be of the order of 20 γ .) Data for the descent portion of



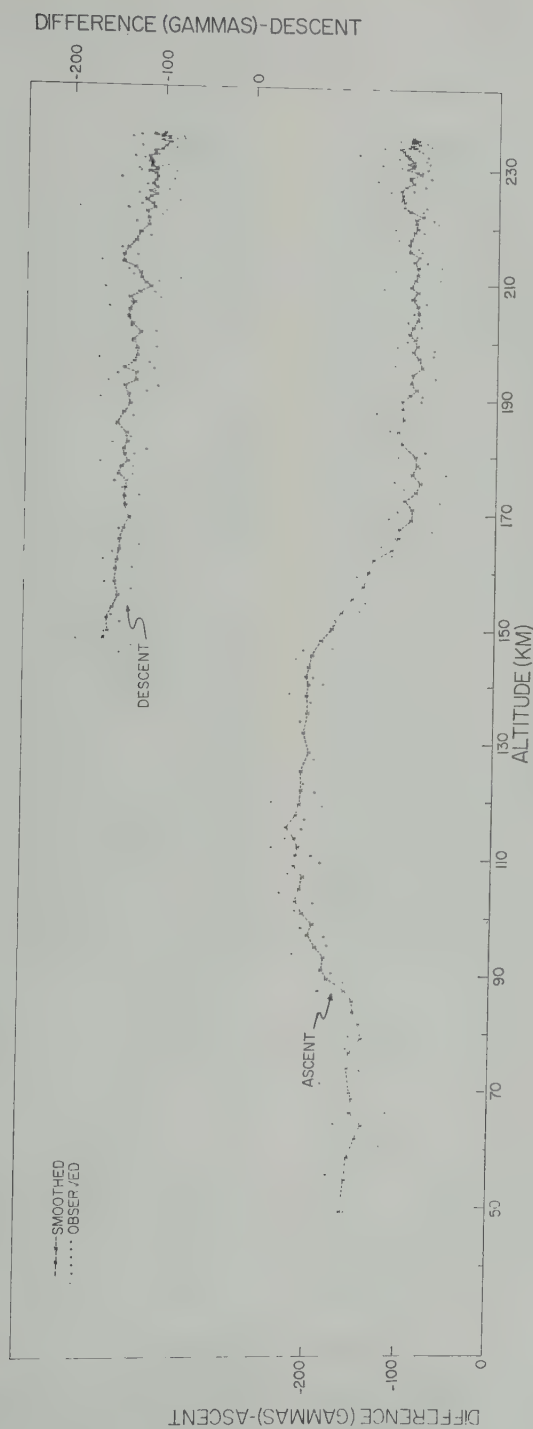


Fig. 4. Observed minus predicted total intensity values versus altitude.

the flight are available only down to 146 km, at which point the tail section was detonated for recovery purposes, thus ending any further telemetry. The difference curve for the descent is essentially a straight line throughout.

It is of interest to note that the Miran position plot (obtained by a triangulation process from the signal of a 600 Mc/s tracking beacon) was blank during just that part of the trajectory in which the anomalous field values were recorded. This suggested the possibility that there might be a causal relationship between the Miran failure and the 'anomalous' field readings, such as might possibly occur in the case of Miran antenna breakdown. However, upon closer inspection of the available films of the Miran signal at the individual receiving sites, it was seen that, whereas the Sugar 1 site near the blockhouse showed a spotty and erratic record during the time in question, the Sacramento Peak station received a continuous strong signal throughout the entire flight.

Thus the anomalous feature of the total intensity curve appears to be a very real effect of conditions that existed at that time in the E region. Attributing this effect to the presence of current systems in the ionosphere would require only a slight extension to the results of *Ratcliffe and Weekes* [1960] wherein the major contribution to atmospheric conductivity in the mid-latitude regions would lie between about 90 and 140 km. However, to postulate a current system for this anomaly is not as straightforward as in the case at the equator where the change in the field due to crossing a uniform current sheet is of magnitude $4\pi j$ cgs emu, where j is the surface current density. Also, if such a current system is in the east-west direction, the effect combines with the total intensity in an essentially scalar fashion. In the present case, however, the large vertical extent of the anomaly and its shape are evidence that an infinite-sheet-current is an inadequate model. Inasmuch as only total intensity measurements were made, this in turn would lead to more ambiguity in the choice of direction for the perturbing field responsible for the anomaly.

As has already been noted, at the time of firing, the ground-based total intensity magnetometer at the launching site showed a decrease of 10 γ from the quiet level of the early morning hours. Comparing this with the magnetogram

from the Tucson Magnetic Observatory, 5° due west of the launching site and at the same latitude, indicates that the observed anomaly was essentially an effect of ΔZ . At the 'equivalent' time of firing at Tucson (10 A.M.), $\Delta Z \approx -16$ γ and $\Delta Y \approx +40$ γ . Using these values with the total intensity curve obtained at the launching site leads to $\Delta Z \approx -16$ γ . These S_e effects observed at the surface can be separated into internal and induced contributions through the relations $\Delta Z_s = \Delta Z_e - \Delta Z_i$ and $\Delta Y_s = \Delta Y_e - \Delta Y_i$. Using the ratio $q = i/e = 0.42$ given by *Rikitake* [1951] leads to the result $\Delta Y_e \approx -\Delta Z_e \approx -27$ γ .

Considering that at this location $I = 61$ γ , therefore $\Delta F(z) \approx 0.875 \Delta Z$ and $\Delta F(y) \approx 0.875 \Delta Y$, it would seem most likely that the observed anomaly in the total intensity was caused by a change in the Z component (that is, the change in F could be caused by $\Delta Z = 90$ γ , $\Delta Y = 760$ γ). Were ΔF a function of ΔZ in the case of crossing an east-west current, the anomaly would be expected to show a reversal of sign, which is not observed.

Assuming then that the observed anomaly is a function of the vertical component, the anomaly mainly to be explained is the sharpness of the change near the boundaries of the E region, and also the dissimilarity between the ascent and descent difference curves. If the observed ΔZ were attributable to the effects of the accepted world pattern of ionospheric currents that accounts for the daily variation, the vertical gradient of ΔZ would be expected to be much less than observed. That the anomaly is so sharply defined in z is a cause of much smaller dimensions. In this case the shape of the anomaly curve could be approximated by the field of a solenoid of order of 10 km in diameter, extending from 100 to 155 km in altitude. Relatively small dimensions in the horizontal plane for the source of the anomaly field would also explain why comparable effects were detected in the descent data down to 146 km. (The horizontal distance covered between the ascent and descent legs of the flight at the 150 km altitude level was 100 km.)

The shape and extent of the observed anomaly suggests a source comparable to the irregular patches of ionization observed as spread F and spread E which have been reported extensively in the literature [*Briggs and Sp*].

Spencer [1955] has further shown that irregularities would take the form of elongated spheroids along the magnetic field lines could be expected to have dimensions in ratio 5 to 1. Radio echoes caused by such columns within the *E* and *F* layers have been reported by Peterson, Villard, Leadabrand, and Gallagher [1955] for the middle latitudes. An irregularity, with a conductivity higher than that of the surrounding medium, could well have a complicated current pattern that would give rise to a magnetic field essentially confined within the source volume.

An explanation suffers, of course, from lack of repeated observations and from the improbability that a single rocket firing would have encountered such a localized phenomenon. However, an alternative interpretation of the observed anomaly in terms of a current sheet at 90 km and a counter-current sheet at 100 km would be still more difficult to justify.

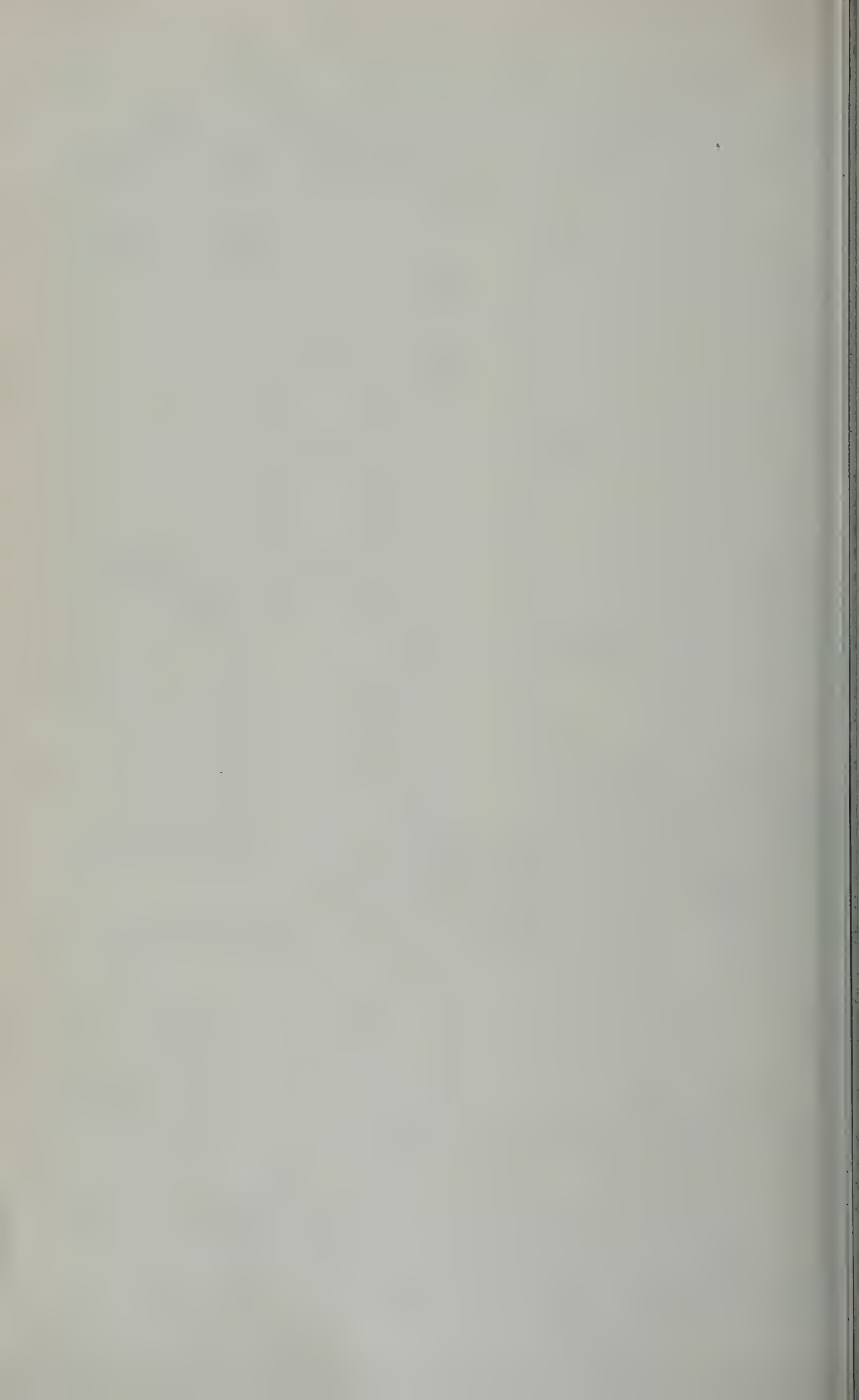
Acknowledgment. We wish to thank Dr. T. T. Shi and Mr. J. McClay for their many helpful discussions on this problem, and also Mr. P. E. for programming the predicted field along the rocket trajectory.

REFERENCES

- A. L., *Tech. Mem.*, Varian Associates, Palo Alto, Calif., 1955.
 B. H., and M. Spencer, Horizontal movements in the ionosphere, *Repts. Progr. in Phys.*, 15, 45-280, 1954.
 L. J., Jr., Magnetic exploration of the upper atmosphere, Doctoral thesis, State University of New York, February 1959.
 S., Rockets and the magnetic exploration of the ionosphere, *Rocket Exploration of the Upper Atmosphere*, edited by R. L. F. Boyd

- and M. J. Seaton, Interscience Publishers, New York, 292-305, 1954.
 Finch, H. F., and B. R. Leaton, The earth's main magnetic field—epoch 1955, *Monthly Notices Roy. Astron. Soc. Geophys. Suppl.*, 7, 314-317, 1957.
 Heppner, J. P., J. D. Stolarik, and L. H. Meredith, The earth's magnetic field above WSPG, New Mexico, from rocket measurements, *J. Geophys. Research*, 63, 277-288, 1958.
 Maple, E., W. A. Bowen, and S. F. Singer, Measurement of the earth's magnetic field at high altitudes at White Sands, New Mexico, *J. Geophys. Research*, 55, 115-126, 1950.
 Marks, S. T., and J. M. Conley, *Mem. Rep. 1214*, Ballistic Research Laboratories, Aberdeen, Md., June 1959.
 Packard, M., Free nuclear precession magnetometer, *Tech. Rept.*, Varian Associates, Palo Alto, California, 1954.
 Packard, M., and R. Varian, Free nuclear induction in the earth's magnetic field, *Phys. Rev.*, 93, 941, 1954.
 Peterson, A. M., O. G. Villard, Jr., R. L. Leadabrand, and P. B. Gallagher, Regularly observable aspect sensitive radio reflections from ionization aligned with the earth's magnetic field and located within the ionospheric layers at middle latitudes, *J. Geophys. Research*, 60, 497-512, 1955.
 Ratcliffe, J. A., and K. Weekes, The ionosphere, *Physics of the Upper Atmosphere*, edited by J. A. Ratcliffe, Academic Press, New York, 377-470, 1960.
 Rikitake, T., *Bull. Earthquake Research Inst.*, 28, 219, 1951.
 Singer, S. F., E. Maple, and W. A. Bowen, Evidence for ionosphere currents from rocket experiments near the geomagnetic equator, *J. Geophys. Research*, 56, 265-281, 1951.
 Spencer, M., The shape of the irregularities in the upper ionosphere, *Proc. Phys. Soc. London B*, 68, 493-503, 1955.

(Manuscript received April 20, 1961; revised June 9, 1961.)



Auroral Effects on the Earth's Electric Field

GEORGE D. FREIER

*School of Physics, University of Minnesota
Minneapolis, Minnesota*

Abstract. On three occasions of auroral appearance near the latitude of Minneapolis (October 7, November 13, and December 1, 1960), measurements of the earth's electric field and time constants of the atmosphere show abnormal conduction currents at the surface of the earth corresponding to 10^4 electrons $\text{cm}^{-2} \text{sec}^{-1}$. A theoretical formulation of fair-weather electric field phenomena is given, and the measurements are interpreted in terms of this theory to show how measurements on the ground are related in general to current generators in the atmosphere.

Introduction. Most people working in meteorology and atmospheric electricity consider the earth's electric field is maintained by storm activity. Thunderstorms act as sources for electrical energy which is stored and dissipated in the region between the ionosphere and the surface of the earth [Wilson, 1939; Whipple [1929] made a study of thunderstorm activity over the surface of the earth and showed that the maximum thunderstorm electric field correlated well with a maximum in fair-weather electric field at about 1800 Greenwich time. The average fair-weather field is about 0.005 st volt cm^{-1} and an average conductivity of $10^{-14} \text{ sec}^{-1}$ gives a fair-weather current density of 5×10^{-7} st amp cm^{-2} . Israel [1955] collected together several results on current measurements which give an average value of 10^{-7} st amp cm^{-2} or a flux of 1.5×10^8 charged particles $\text{cm}^{-2} \text{sec}^{-1}$.

Whipple and Bartels [1940] discussed the relation of the earth's magnetic field with the flux of particles from the sun and showed that the two charge components can be separated in complicated ways so that some enter the atmosphere and there dissipate their energy. Scherrer [1955] has worked out the trajectories of individual charged particles and showed how they might follow the earth's magnetic field to produce the aurora. More recently [1958] and Arnoldy, Hoffman, and [1960] have found, at levels attained in magnetic storms, which they ascribe to the stopping in the upper atmosphere

after leaving the Van Allen belt. Anderson and Enemark [1960] have found the flux of electrons with energies above 25 kev to be $4 \times 10^6 \text{ cm}^{-2} \text{sec}^{-1}$. If this high flux were predominantly of one sign of charge, an aurora in fair weather should lead to an increased fair-weather conduction current in the atmosphere. The auroral electrons lose their energy at altitudes of approximately 130 km and create space charge. A fraction of the electric field lines should terminate on the earth and lead to increased conduction currents in the lower atmosphere.

Theory. Observations of these effects on the ground are somewhat complicated because the field lines have to penetrate a conducting atmosphere, but the response at the ground may also tell us more about the properties of the atmosphere itself.

The situation can be analyzed as a boundary value problem. It is well known that the atmosphere can carry space charge ρ ; that it has a finite conductivity λ ; that it can support an electric field \mathbf{E} ; that there is a conduction current \mathbf{j} given by $\mathbf{j} = \lambda \mathbf{E}$; and that there may be convection current densities \mathbf{J} carrying charge from place to place. For each region of the atmosphere, then, the following three equations can be written:

$$\text{div } \mathbf{E} = \frac{4\pi}{\epsilon} \rho$$

$$\frac{\partial \rho}{\partial t} + \text{div } \mathbf{j} = -\text{div } \mathbf{J}$$

$$\mathbf{j} = \lambda \mathbf{E}$$

The first equation tells how \mathbf{E} changes in passing through a charge distribution ρ ; the second is the continuity equation in the case that some charge is being distributed by convection currents; and the last is Ohm's law for each element of space.

These simultaneous equations can be solved by differentiating the first equation with respect to time, and combining it with the other two equations to give

$$\text{div} \left[\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{\epsilon} (\lambda \mathbf{E} + \mathbf{J}) \right] = 0$$

The volume integral of this over the atmosphere is also a surface integral over the boundaries of the atmosphere, one boundary being the earth and the other being the top of the atmosphere. When this is done we have

$$\oint_0 \left[\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{\epsilon} (\lambda \mathbf{E} + \mathbf{J}) \right] \cdot d\mathbf{S} \\ = \oint_h \left[\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{\epsilon} (\lambda \mathbf{E} + \mathbf{J}) \right] \cdot d\mathbf{S}$$

Here the integral on the left is at ground level and the integral on the right is at a height h in the atmosphere. The surface normal is directed inward in the right-hand term, so that it will be positive for charge coming to the atmosphere from outer space.

If we accept the possibility of charge entering from outer space and think of it as a cloud with a net charge moving toward the earth, it will make itself known at the upper surface first as the displacement current density

$$\frac{\epsilon}{4\pi} \frac{\partial \mathbf{E}}{\partial t}$$

and then later as a conduction current $\mathbf{j} = \lambda \mathbf{E}$ or possibly as a convection current \mathbf{J} . We cannot determine which form the current has, so we will simply collect the terms into a single term \mathbf{J} at the top of the atmosphere.

At the bottom of the atmosphere, the field is normal to the conducting earth and induces a surface charge density $\sigma_0 = E_0/4\pi$. The conductivity in the lower layer of the atmosphere may be written as

$$\lambda_0 = \epsilon_0/4\pi\tau_0$$

where τ_0 is the relaxation time of the air. When

we make these substitutions, our integral over the boundary becomes

$$\oint_0 \left[\frac{\partial \sigma_0}{\partial t} + \frac{\sigma_0}{\tau_0} + \mathbf{J}_0 \right] \cdot d\mathbf{S} = \oint_h \mathbf{J} \cdot d\mathbf{S}$$

The dielectric constant ϵ_0 is assumed everywhere equal to unity.

During ideal fair weather, measurements show that variations in the field quantities tend to be vertical in that measurements of $d\sigma_0/dt$ at stations separated by several miles produce similar records. If we then outlined a spherical pillbox element in the atmosphere, including a region of fair weather, we could neglect contributions to the surface integral along the vertical sides of the element and be concerned only with the ends, one at the surface of the earth and the other at a great height. These two areas would be approximately equal, so that we can write an equation for the average value of the integrals as

$$\frac{d\sigma_0}{dt} + \frac{\sigma_0}{\tau_0} = \mathbf{J} - \mathbf{J}_0$$

Measurements of σ_0 and τ_0 at the ground as a function of time, should give information about $\mathbf{J} - \mathbf{J}_0$. If we observe variations in the measured values of $d\sigma_0/dt + \sigma_0/\tau_0$ that are characteristic of fair-weather fields and that coincide well with visually observed phenomena that might possibly be bringing charge in from an upper boundary, we would be inclined to believe that \mathbf{J} might be responsible for the variations. \mathbf{J}_0 , the convection current density from the atmosphere to the earth across the lower boundary, is not measured directly in this experiment but there is no reason to believe that it should suddenly vary coincidentally with auroral activity. During nights of fair weather \mathbf{J}_0 is found to be exceedingly constant.

The theory developed here shows that increased ionization of the air, either direct by auroral particles or by X rays produced by cosmic rays, should not produce a variation in the average values of $d\sigma_0/dt + \sigma_0/\tau_0$. In fact, any processes within the region enclosed by our boundary will not affect the values on the boundary unless they are highly localized and then vitiate the assumption of only vertical variation. Processes that transport charge across the boundaries should make $d\sigma_0/dt + \sigma_0/\tau_0$ vary

ing fair weather where σ_0 is approximately constant, J_0 is thought to be zero, and $J = J_f$, J_f is the charge re-entering the atmosphere being deposited in the conducting ionosphere by thunderstorm activity at other regions of the earth. The boundary equation then reduces to $\sigma_0/\tau_0 = J$ for $E_0 = \lambda_0 J_f$. These are established relations.

In the presence of considerable magnetic activity, we should include a contribution to \mathbf{E} according to the curl $\mathbf{E} = -(1/c) (\partial \mathbf{B}/\partial t)$. This was neglected at the outset because the observed field had the same sign throughout the observations which would imply that \mathbf{B} must be changing directionally with time for times of the order of hours. This would lead to values of \mathbf{B} much larger than those observed. Furthermore, we neglect for these phenomena the relation that

$$\int \frac{\partial \mathbf{B}}{\partial t} dt = 0$$

we should expect \mathbf{E} to reverse sign during the course of observation. As there was no indication that this happened, the possible origin of \mathbf{E} according to Faraday's law of electromagnetic induction is neglected.

Experimentation. An electric field mill constructed with proper time constants and properly spaced will give a measurement of σ_0 . Our instrument was on a small tripod placed among other field projections of about the same size. Corrections on a scale model in a conduction experiment indicated that corrections of less than 75 per cent should be made for geometrical factors affecting the field. This geometry factor could be as much as a 75 per cent error in the measured values of σ_0 , but relative values are correct to ± 5 per cent.

An isolated conductor having a charge on it suspended in the air will lose its charge with a characteristic time constant of the air, τ_0 , since the air is the only leakage path to ground. Resistance and capacity have the same geometrical factors for a conductor of any shape. The charged conductor will also have a characteristic electric field that will decay with the time constant, τ_0 . If one insulates the stator plates of the mill very well and charges them, the mill can measure its own field in a shielded cage. The cage can easily shield against external fields and still allow for a free exchange of air.

A small air condenser was used to couple the mill to the amplifiers, and air was circulated through this condenser. The working equation for the device is $E' = E_0 e^{-t/\tau_0}$. The instrument has a resistance at least of 10^{10} ohms, will readily detect small amounts of radioactivity producing ionization in its proximity, and will maintain its high resistance in dense clouds of steam.

With a regular mill measuring σ_0 and the second mill measuring τ_0 information about J can be obtained. One combination like this giving continuous records was placed on top of the Physics Building within the city of Minneapolis. At the time of these measurements, a field mill giving σ_0 was operated at a second station at a small airport 10 miles north of the city. The records from the city station are labeled A; those from the airport station are labeled B.

The earth's magnetic field in Minneapolis is being recorded by J. R. Winckler and his group, and sections of this record are included to indicate the extent of the auroral activity.

Attempts to photograph the aurora failed, but the aurora that produced the record of October 7 would be classified by Störmer [1955] as auroral corona (C) in which rays appear near the zenith and seem to be directed toward one point. The strong coronal aurora was visible only from 0330 to 0400 UT, and it correlates timewise very well with the dip in the electric field and magnetic field. The auroral display which produced the record on November 13 was more difficult to classify but would in general be of type II in Störmer's classification. The display of December 1 was quite similar to the November 13 display but less intense. On each of these occasions the sky was absolutely clear of clouds at the time of these records.

The records shown in Figures 1a, b, and c have been smoothed, and the length of the vertical bars indicates the amplitude of the variations about the smooth line drawn through the points.

Results. The beginning of each record was chosen to match the fair-weather field. The values leading to the calculation of the fair-weather current density are listed in Table 1. Since τ_0 was not measured at station B, it was calculated by making J_f the same at both stations. Note that the fields are much higher than

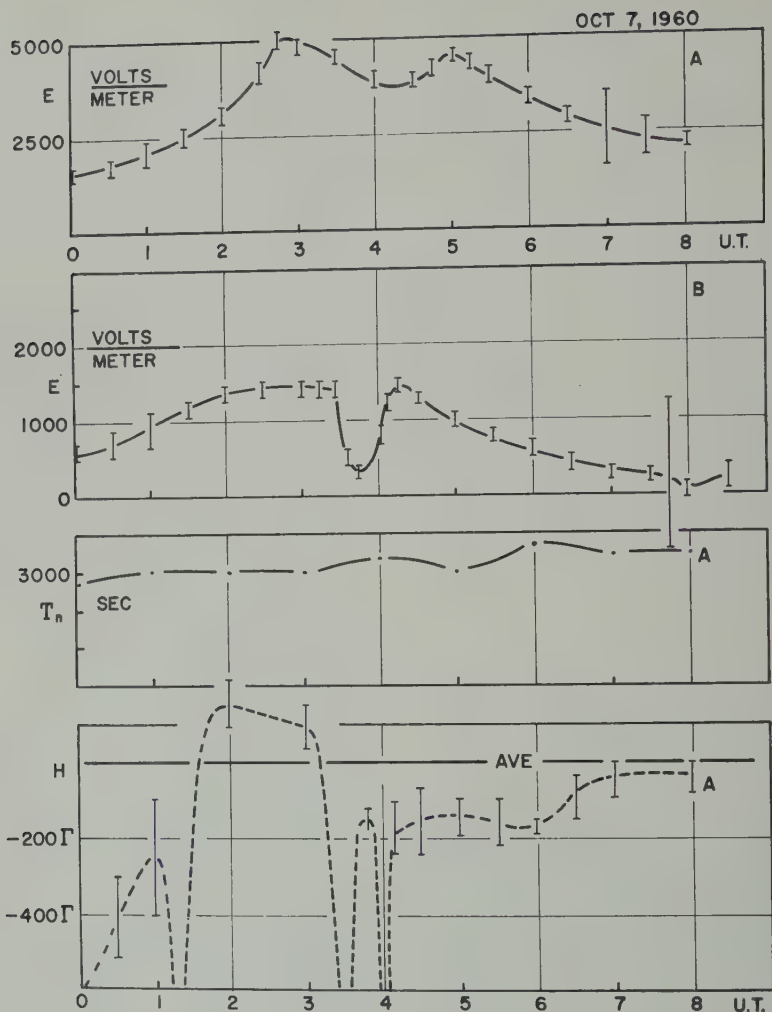


Fig. 1a.

the normally quoted night values and that τ_0 also much exceeds the quoted values. Thus, rather large currents are indicated. There could be some error in the measurement of these long time constants, where any leakage in the instrument would tend to make the values of τ_0 too small. The combined errors in absolute values of σ_0 and τ_0 are probably responsible for the currents' being above average. Geometrical factors render σ_0 too large at station A and too small at station B, but no corrections were included in Tables 1 and 2.

On October 7 and November 13 the field rises before taking rather sharp dips. The magnetic field was quite variable during this period, but

there was very little if any visible aurora to correlate with the rise. On October 7 there had been considerable burning of leaves and during the early evening, which would lead to a current generator giving J_0 a value. Separate measurements on city combustion products show that J_0 is directed upward. Normally the fair-weather field is uniform or becomes smaller at this time of day. The rise could be a contribution due either to J_0 or to J .

The dips in the electric field would correspond to electrons directed toward the station. Table 2 shows the estimated changes in the electric field and the flux of electrons which could produce this charge. Considering all the ap-

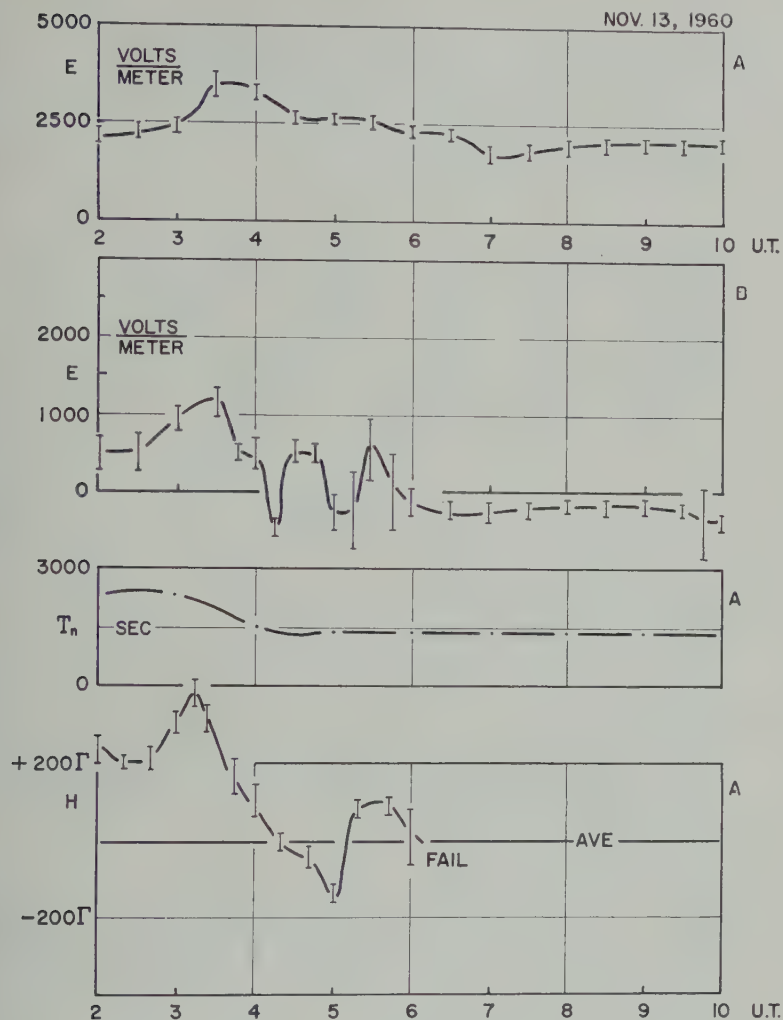


Fig. 1b.

s and errors in measurement, it seems 10^4 particles $\text{cm}^{-2} \text{sec}^{-1}$ is a best value for current at the surface of the earth due to aurora. Note that on November 12 station maintained a negative fair-weather field. It must be remembered that this value of the density is a net value for both positive and negative particles that may be incident at top of the atmosphere. Also, the current is measured in this experiment represent the fraction of the current that has been carried through a path of highest resistance. The factors were properly taken into account. It seems that this measurement of 10^4

particles $\text{cm}^{-2} \text{sec}^{-1}$ is not inconsistent with Anderson's value of 4×10^6 particles $\text{cm}^{-2} \text{sec}^{-1}$. Although periods of auroral activity near the latitude of Minneapolis are infrequent, apparently this can be a source of some of the charge on the earth.

Near the auroral zones a crackling noise associated with auroras is sometimes reported. Chapman and Bartels [1940] have summarized the situation, concluding that the noise exists. Could this be due to electrical breakdown? The usual relation for the breakdown field is $E_b/P = 42 \text{ volts cm}^{-1} (\text{mm of Hg})^{-1}$. If we assume that the pressure of the atmosphere

DEC 1, 1960

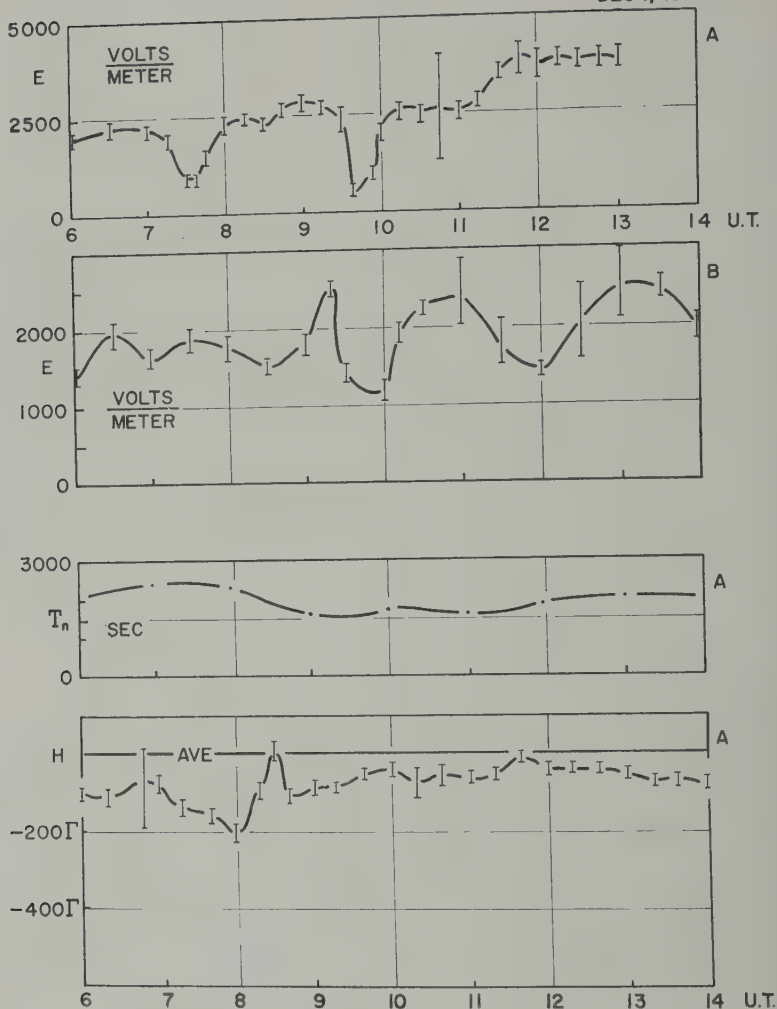


Fig. 1c.

Fig. 1. Smoothed records of the earth's electric field at two stations, the time constant of the air, and the earth's magnetic field. The letter A implies measurements at a station within the city, and B implies measurements outside the city. The vertical bars on the field records show magnitudes of fluctuation about the smoothed average. Greenwich time is plotted on the abscissa.

varies with the altitude, h , as $\epsilon^{-h/H}$, then in cgs units $E_B = 100^{-h/H}$ st volts cm^{-1} . write the equation as

The equation

$$\text{div} \left[\frac{\partial \mathbf{E}}{\partial t} + \frac{4\pi}{\epsilon} (\lambda \mathbf{E} + \mathbf{J}) \right] = 0$$

may be applied to any region of the atmosphere.

Consider the time integral of the quantity in brackets. If we let $1/\tau = 4\pi\lambda/\epsilon$, we may

$$\frac{\partial \mathbf{E}}{\partial t} + \frac{\mathbf{E}}{\tau} = \frac{4\pi \mathbf{J}_a}{\epsilon_0}$$

where \mathbf{J}_a is a current density due to the through a top layer and

$$\frac{\epsilon_0}{4\pi} \left[\frac{\partial \mathbf{E}}{\partial t} + \frac{\mathbf{E}}{\tau} \right]$$

TABLE 1. Values of the Fair-Weather Field and Time Constants for the Air with the Corresponding Calculated Values of the Fair-Weather Currents at the Two Stations on Days of the Aurora

order to obtain time constants at station B, it was assumed that the fair-weather current was the at both stations. No corrections are made for geometry factors.

	Time	Station	E_0 , volt m^{-1}	σ_0 , esu cm^{-2}	τ_0 , sec	J_f , st amp cm^{-2}
3	0000	A	1500	4.0×10^{-3}	2.8×10^3	14×10^{-7}
		B	500	1.3×10^{-3}	0.9×10^3	14×10^{-7}
	0200	A	2200	5.8×10^{-3}	2.3×10^3	25×10^{-7}
		B	500	1.3×10^{-3}	0.5×10^3	25×10^{-7}
	0600	A	2000	5.3×10^{-3}	2.1×10^3	25×10^{-7}
		B	1400	3.7×10^{-3}	1.5×10^3	25×10^{-7}

total current through the bottom layer placed at an arbitrary height in the atmosphere. In this case, as we neglect local convection currents, E is caused only by J_a . The equation may be integrated to give

$$E = \epsilon^{-\sigma_0 t/\tau} \int_0^t \frac{4\pi}{\epsilon_0} J_a \epsilon^{\sigma_0 t'/\tau} dt'$$

is constant in time at the height under consideration, and if J_a was zero until $t = 0$ then took on a constant value J_a , for the time of time t , the equation for E becomes

$$E = \frac{4\pi}{\epsilon_0} J_a \tau [1 - \epsilon^{-t/\tau}]$$

Measurement of atmospheric conductivity, relaxation time of the air is found to vary approximately as $\tau = \tau_0 \epsilon^{-h/H}$. For electrical breakdown in the atmosphere, we then have condition that $E > E_B$, or

$$\frac{\pi}{\epsilon_0} J_a \tau_0 \epsilon^{-h/H} [1 - \epsilon^{-t/\tau}] > 100 \epsilon^{-h/H}$$

quantity in brackets will be unity if $t > \tau$, and if we use $\tau_0 = 10^3$ sec the inequality becomes

ity becomes

$$J_a > 8 \times 10^{-3} \text{ st amp } cm^{-2}$$

No currents as large as this were detected in this experiment, but this value is quite close to Anderson's flux. On October 7 between 0700 and 0800 the records had spikes as large as 1500 volts m^{-1} with rise times as short as 30 sec. If we use Gish's [1944] equation for conductivity as a function of altitude, we obtain $\tau < 30$ sec for altitude above 4 km. Any noise caused by breakdown would have to be produced at altitudes above 4 km if the above field changes were responsible for them. The recording equipment in this experiment was not fast enough or sensitive enough to follow some of the faster light flashes that occur during auroras; but, if the bright flashes are caused by corresponding high fluxes of electrons which in turn produce breakdown fields, our condition $t > \tau$ can be satisfied only at a very great altitude. Increased ionization produced by X rays from the stopping electrons would make τ smaller at any given point and allow breakdown to occur more readily at lower altitudes. Larger fluxes than

TABLE 2. Values of Electrical Field Changes Due to Auroral Currents along with the Corresponding Calculated Current Densities and Electron Fluxes

Station	ΔE , volt m^{-1}	τ_0 , sec	J_a , st amp cm^{-2}	Flux No. $cm^{-2} \text{ sec}^{-1}$
A	1500	3×10^3	13×10^{-7}	2.8×10^3
B	1500	0.9×10^3	44×10^{-7}	9.2×10^3
A	1500	1.5×10^3	27×10^{-7}	5.6×10^3
B	1500	5×10^3	80×10^{-7}	16.6×10^3
A	3000	2×10^3	40×10^{-7}	8.4×10^3
B	1000	1.5×10^3	18×10^{-7}	3.8×10^3

this limiting flux will also allow breakdown to occur at lower altitudes.

Other attempts to make electric field measurements on auroras at this locality have failed, mostly because of the presence of clouds. The data reported in this paper are the only records obtained to date where the sky was perfectly clear and there was little or no wind. These stations are located at the foot of magnetic lines of force that pass through the center of the outer Van Allen belts. An aurora in this locality always produces X rays, whereas an aurora farther north may produce X rays on some occasions and not on others. This may imply that we are observing charging effects due to the incidence of higher-energy particles which originally were in a neutral plasma but which can suffer a greater range separation on coming to rest in the upper atmosphere.

We have inferred a uniform divergence on the current at ground level from a measurement at only two stations. This divergence seems to be similar on three separate occasions and coincides with the auroral activity that might be producing a divergence of current into the top of the atmosphere. In spite of the limited data, it seems that the only conclusion to draw is that the aurora brought a net negative charge into the atmosphere.

Acknowledgments. I wish to thank Professor J. R. Winckler and his group for making magnetic field records available and for information about auroras in general. I also wish to thank Professor

S. Chapman and Professor H. Mantis for discussions in formulating the problem.

This work was supported in part by a grant from the National Science Foundation.

REFERENCES

- Anderson, K. A., Soft radiation events at high altitudes during a magnetic storm, *Phys. Rev. Lett.*, **111**, 1397-1405, 1958.
- Anderson, K. A., and D. C. Enemark, I. observations of X rays in the auroral region, *J. Geophys. Research*, **65**, 3521-3538, 1960.
- Arnoldy, R., R. Hoffman, and J. R. Winckler, Observations of Van Allen radiation during geomagnetic storms, *Proc. COE First International Space Science Symposium*, Nice, 1960.
- Chapman, S., and J. Bartels, *Geomagnetism and Geoelectricity*, vol. 1, pp. 1-542, vol. 2, pp. 543-1049, Clarendon Press, Oxford, 1940.
- Gish, O. H., Evaluation and interpretation of the columnar resistance of the atmosphere, *Terrestrial Magnetism and Atmospheric Electricity*, **159**, 159-168, 1944.
- Israel, H., Der Luftelektrische Vertikalstrom, *Ann. Geophys.*, **10**, 93-116, 1954.
- Störmer, C., *The Polar Aurora*, Oxford University Press, New York, 392 pp., 1955.
- Whipple, F. J. W., On the association of the diurnal variation of electrical potential gradient in fine weather with distribution of the magnetic storms on the globe, *Quart. J. Roy. Meteor. Soc.*, **55**, 1-17, 1929.
- Wilson, C. T. R., A theory of thundercloud electrification, *Proc. Roy. Soc. London, A*, **236**, 211-221, 1956.

(Manuscript received January 26, 1961; revised June 12, 1961.)

The Spectrum and the Total Intensity of Electromagnetic Waves Scattered from an Ionized Gas in Thermal Equilibrium in the Presence of a Static Quasi-Uniform Magnetic Field¹

J. RENAU, H. CAMNITZ, AND W. FLOOD

*Cornell Aeronautical Laboratory, Inc.
Buffalo, New York*

Abstract. The spectrum of electromagnetic waves scattered from thermal electron density fluctuations in an ionized gas in the presence of a static, quasi-uniform, magnetic field has been derived through the use of the Nyquist noise theorem. General analytical results of the spectrum are given and graphs presented for the special case of backscattering.

INTRODUCTION

In a harmonic electromagnetic plane wave propagates through an ionized gas in thermal equilibrium, it will be scattered by the ever-present thermal fluctuations in electron number density. The theory of this scattering in the presence of external static fields has recently been treated by a number of authors [Dougherty and Farley, 1960; Fejer, 1960; Renau, 1960; Renau, 1960a, 1960b] and confirmed experimentally by Bowles [1959] and Pineo, Kraft, and Renau [1960]. These studies clearly highlight the important role of the positive ions of the gas and their effect on the electrons through the ion-ion interaction. Laaspere [1960] considered the effects of an external magnetic field on scattering phenomena using an independent electron model for the gaseous system. The work of Laaspere therefore treats the problem thoroughly in the region of wavelengths less than the Debye length.

In this paper we derive the equations of electromagnetic scattering from a nondegenerate, ionized gas in thermal equilibrium in the presence of a uniform magnetic field for wavelengths of the order of the Debye length. The treatment is similar to that employed by Farley and Fejer [1960] in their discussion of the magnetic field-free case. The generalized Nyquist theorem is employed to obtain the

relation between the power spectrum of the electron density fluctuations and the generalized admittance tensor of the system. This generalized admittance tensor, which includes the effects of charge interactions and the external magnetic field, is expressed in terms of the admittance tensors for single species of noninteracting particles in thermal equilibrium. These considerations lead to the formal solution of the scattering problem.

The formal mathematical derivations can be found in sections 1, 2, and 4 and Appendices A, B, and C. The analytical expressions and the graphical representations of the backscatter spectrum are given in section 3, and the results are summarized in section 5.

COMMONLY USED SYMBOLS

ω_0 = the incident wave angular frequency.

ω' = the observed angular frequency of scattered waves.

$\omega = \omega' - \omega_0$.

$\mathbf{q} = \mathbf{k}_{\text{inc}} - \mathbf{k}_s$ with \mathbf{k}_{inc} the incident wave propagation vector and \mathbf{k}_s the propagation vector corresponding to the direction from scatterer to observer.

$|\mathbf{q}| = 2\omega_0/c \sin \epsilon/2 = q$.

ϵ = angle between \mathbf{k}_{inc} and \mathbf{k}_s .

c = velocity of light in free space.

$d_0 \equiv (KT/4\pi\bar{N}e^2)^{1/2}$ = the Debye length in cgs units.

K = Boltzmann constant.

T = Kelvin temperature.

$\bar{N}^i = \bar{N}^e \equiv \bar{N}$ defined as mean ion number density equal to mean electron number density.

A preliminary summary of this paper was presented at the General Assembly of the International Union of Pure and Applied Physics (IUPAP) in Boulder, Colorado, August 11-14, 1960. At that meeting Drs. J. H. Dumbarton and Hagfors also presented their independent results on this topic.

e = unit of charge.

$\theta^e \equiv (\omega/q) (m^e/2KT)^{1/2}$ with m^e as the mass of an electron.

$\theta^i \equiv (\omega/q) (m^i/2KT)^{1/2}$ with m^i as the mass of an ion.

$\phi^e \equiv (\Omega^e/q) (m^e/2KT)^{1/2}$ and for the case of backscatter equals $\lambda/4\pi R^e$.

$\phi^i \equiv (\Omega^i/q) (m^i/2KT)^{1/2}$ and for the case of backscatter equals $\lambda/4\pi R^i$.

$R^e \equiv (2KT/m^e)^{1/2} 1/\Omega^e$ referred to as the electron gyroradius.

$R^i \equiv (2KT/m^i)^{1/2} 1/\Omega^i$ referred to as the ion gyroradius.

$\Omega^e \equiv e|B_0|/m^e c$ = angular gyrofrequency of an electron.

$\Omega^i \equiv e|B_0|/m^i c$ = angular gyrofrequency of an ion.

B_0 = externally applied uniform magnetic field.

δ = the acute angle between \mathbf{q} and \mathbf{B}_0 . In the case of backscatter this is the acute angle between the incident plane wave direction and the magnetic field.

1. THE RELATION BETWEEN THE SPECTRAL SCATTERING CROSS SECTION AND THE POWER SPECTRUM OF THE ELECTRON NUMBER DENSITY FLUCTUATIONS

The spectral differential scattering cross section per unit volume is given by

$$\sigma(\omega_0 \pm \omega) = \frac{V}{2} r_e^2 \sin^2 \xi \overline{|\Delta N(\mathbf{q}, \omega)|^2} \quad (1)$$

where

r_e = classical electron radius = 2.8×10^{-13} cm.

ξ = the angle between the incident electric field intensity and \mathbf{k}_s .

V = the scattering volume.

$\overline{|\Delta N(\mathbf{q}, \omega)|^2}$ = the power spectrum of

$$\Delta N_a(t) = \frac{1}{V} \int_V \Delta N(\mathbf{r}, t) \exp(j\mathbf{q} \cdot \mathbf{r}) dV \quad (2)$$

where $\Delta N(\mathbf{r}, t)$ denotes the stochastic electron density fluctuations in the gas.

Equation 1 is derived in Appendix A under the following conditions: (1) single scattering is a good approximation of the scattering process; (2) the operating frequency is much greater than the electron plasma frequency, the electron collision frequency, and the electron gyrofrequency (so that the dielectric constant of the

gas is substantially unity); (3) $\Delta N_a(t)$ is stationary relative to $e^{j\omega_0 t}$; (4) the gas is free of external static fields except for an external magnetic field; (5) the gas is in a state of macroscopic equilibrium. Condition 5 can be replaced by the weaker condition of quasi-macroscopic equilibrium for the range of scale sizes λ so that equation 1 should retain an approximate validity for the case of scattering from large-scale, inhomogeneous, isotropic, turbulent irregularities.

To obtain the differential cross section per unit volume, integrate (1) with respect to ω from $-\omega_0$ to $+\infty$, i.e.,

$$\begin{aligned} \sigma_t &= \int_0^\infty \sigma(\omega') d\omega' \\ &= \frac{V}{2} r_e^2 \sin^2 \xi \int_{-\omega_0}^\infty \overline{|\Delta N(\mathbf{q}, \omega)|^2} d\omega \\ &\simeq \frac{V}{2} r_e^2 \sin^2 \xi \int_{-\infty}^\infty \overline{|\Delta N(\mathbf{q}, \omega)|^2} d\omega \end{aligned}$$

The spectrum and the total intensity of electromagnetic scattering will be known when an explicit expression is obtained for $\overline{|\Delta N(\mathbf{q}, \omega)|^2}$ in terms of the parameters characteristic of the physical state of the gas. In this paper we consider the gas to be in a state of thermal equilibrium.

2. DERIVATION OF THE RELATION BETWEEN THE POWER SPECTRUM OF THE ELECTRON NUMBER DENSITY FLUCTUATIONS OF SCALE \mathbf{q} AND THE EFFECTIVE ADMITTANCE TENSOR FOR ELECTRON FLOW

The expression for the power spectrum of electron density fluctuations of scale \mathbf{q} is related to the admittance tensor for electron flux in the system by an application of the generalized Nyquist theorem. This theorem has been widely discussed in the literature (see references at the end of the paper). Though shown by Richardson [1955], Nyquist's theorem fails to be valid for nonequilibrium systems, it provides a powerful method for linear systems in thermal equilibrium. For convenience of presentation a brief statement of the Nyquist theorem will be given in a form readily applicable to the problem at hand.

Consider a gaseous system contained within a cubical volume of side L . All field quantities within this system can be expanded in series

exp $(-j\mathbf{k}_\mu \cdot \mathbf{r})$, where

$$\exp [j(\mathbf{k}_\mu - \mathbf{k}_{\mu'}) \cdot \mathbf{r}] dV = V \delta_{\mu\mu'}$$

region of integration is the entire system. Suppose that the system is linear, so the application of a force field

$$\mathbf{F}(\mathbf{r}, t) = \mathbf{F}_0 \exp [j(\omega t - \mathbf{k}_\mu \cdot \mathbf{r})]$$

species of the system produces a forced electron flux $\mathbf{I}_0 \exp [j(\omega t - \mathbf{k}_\mu \cdot \mathbf{r})]$ in a field quantity where

$$\mathbf{I}_0 = \mathbf{Y}(\mathbf{k}_\mu, \omega) \cdot \mathbf{F}_0$$

\mathbf{Y} denotes the tensor admittance for the amplitude \mathbf{I}_0 under the application of the amplitude \mathbf{F}_0 . Further, suppose that the average power delivered to the system by the force $\mathbf{F}(\mathbf{r}, t)$ is given by $(V/2)\text{Real}(\mathbf{F}_0^* \cdot \mathbf{I}_0)$. Poynting's theorem states that, when this system is in thermal equilibrium, the field quantity

$$\Delta \mathbf{I}(\mathbf{r}, t) = \mathbf{I}(\mathbf{r}, t) - \overline{\mathbf{I}(\mathbf{r}, t)}$$

is a stochastic function with the power spectra Fourier amplitudes

$$\Delta \mathbf{I}(\mathbf{r}, t) = \frac{1}{V} \int_V \Delta \mathbf{I}_i(\mathbf{r}, t) \exp (j\mathbf{k}_\mu \cdot \mathbf{r}) dV$$

orthogonal components, $\Delta \mathbf{I}_i(\mathbf{r}, t)$, given by

$$\begin{aligned} & \overline{|\Delta \mathbf{I}_i(\mathbf{r}, t)|^2} d\omega \\ &= \frac{2KT}{\pi V} \text{Real } Y_{ii}(\mathbf{k}_\mu, \omega) d\omega \end{aligned} \quad (4)$$

are ready to apply this theorem to the system of interest. This system is an electrically homogeneous, ionized gas consisting of a single species of singly ionized electrons, and neutral particles, in thermal equilibrium in the presence of a uniform magnetic field. In applying the theorem, the route is followed by following the procedure of *Dougherty* [1960]. Suppose for the moment that

$$\mathbf{F}^e = \mathbf{F}_0^e \exp [j(\omega t - \mathbf{k}_\mu \cdot \mathbf{r})]$$

is applied to the electrons alone, and the electron flux calculated neglecting interactions between charged particles. If the applied

applications of this theory, the extension to a uniform, static magnetic field is evident.

force is weak, the system is linear, so that the resulting forced-response electron flux is given by

$$\mathbf{S}^e = \mathbf{Y}^e \cdot \mathbf{F}^e \quad (5)$$

where \mathbf{Y}^e denotes the admittance tensor for electrons alone. Similarly, for a weak force \mathbf{F}^i applied to the positive ions alone, an admittance tensor \mathbf{Y}^i can be defined for ions alone. By utilizing the single-species tensors \mathbf{Y}^e and \mathbf{Y}^i , the effect of the perturbation in the electromagnetic particle-particle interactions due to the applied force field \mathbf{F} can be included, as will be shown below.

In formulating the equations to determine the spectrum of the electron density fluctuations, only the particular propagation vector $\mathbf{k}_\mu = \mathbf{q}$ will be chosen. If the spectrum of the electron flux of scale \mathbf{q} is obtained, the corresponding spectrum of the electron density fluctuations of scale \mathbf{q} can be obtained by application of the continuity equation. In applying the Nyquist theorem to obtain the spectrum of the electron flux, the applied force field can do work on the electrons only. It is assumed that a force field

$$\mathbf{F} = \mathbf{F}_0 \exp [j(\omega t - \mathbf{q} \cdot \mathbf{r})]$$

is applied to the system, and this force field only acts on the electrons. In perturbing the particle fluxes and densities, the force \mathbf{F} will perturb the Coulomb interaction fields between particles, giving rise to a forced-response electric field of the form $\mathbf{E} = \mathbf{E}_0 \exp [j(\omega t - \mathbf{q} \cdot \mathbf{r})]$. Thus, a net force field $\mathbf{F} - e\mathbf{E}$ will be applied to the electrons, and a net force field $e\mathbf{E}$ will be applied to the positive ions.³ Letting \mathbf{S}^e and \mathbf{S}^i denote the forced-response electron and positive ion fluxes, respectively, it follows that

$$\mathbf{S}^e = \mathbf{Y}^e \cdot (\mathbf{F} - e\mathbf{E})$$

$$\mathbf{S}^i = \mathbf{Y}^i \cdot (e\mathbf{E})$$

These particle fluxes give rise to a net current density

$$\mathbf{J} = e(\mathbf{S}^i - \mathbf{S}^e)$$

³ There is a forced-response magnetic field, \mathbf{B} , which accompanies the forced-response electric field \mathbf{E} . The average force per electron due to \mathbf{B} is $(e/c)\mathbf{u} \times \mathbf{B}$, where \mathbf{u} is obtained by averaging \mathbf{v} over the velocity distribution. For a linear response, $\mathbf{u} \times \mathbf{B}$ must be evaluated only to first order in \mathbf{F} . But $\mathbf{u} \times \mathbf{B} = 0$ in first order since $\mathbf{u} = 0$ in zero order.

which is the source of the response field \mathbf{E} . To obtain the equation for \mathbf{E} we eliminate the magnetic field, \mathbf{B} , from Maxwell's equations and obtain the wave equation

$$\nabla \times \nabla \times \mathbf{E} + \frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = -\frac{4\pi}{c^2} \frac{\partial \mathbf{J}}{\partial t} \tag{6}$$

Since \mathbf{J} has the form $\mathbf{J}_0 \exp [j(\omega t - \mathbf{q} \cdot \mathbf{r})]$, the desired solution of (6) is

$$\mathbf{E} = \mathbf{G}^{-1} \cdot \mathbf{J}$$

where \mathbf{G}^{-1} denotes the inverse of

$$\mathbf{G} = \frac{j c^2}{4 \pi \omega} \left[\left(q^2 - \frac{\omega^2}{c^2} \right) \mathbf{I} - \mathbf{q} \mathbf{q} \right] \tag{7}$$

In (7), \mathbf{I} denotes the identity tensor. The above system of equations can now be solved for the response \mathbf{S}^e , yielding

$$\mathbf{S}^e = \mathbf{Y}' \cdot \mathbf{F} \tag{8}$$

where \mathbf{Y}' , the effective admittance tensor for the electron flux, is given by

$$\mathbf{Y}' = \left(\mathbf{Y}^i - \frac{\mathbf{G}}{e^2} \right) \cdot \left(\mathbf{Y}^i + \mathbf{Y}^e - \frac{\mathbf{G}}{e^2} \right)^{-1} \cdot \mathbf{Y}^e \tag{9}$$

Since the mode functions $\exp (-j \mathbf{k}_\mu \cdot \mathbf{r})$ are orthogonal on V , and the force field \mathbf{F} acts only on the electrons, the average power delivered to the system by the force \mathbf{F} is

$$\frac{1}{2} \text{Real} \int_V \mathbf{F}^* \cdot \mathbf{S}^e dV = \frac{V}{2} \text{Real} (\mathbf{F}^* \cdot \mathbf{S}^e)$$

Therefore, the application of Nyquist's theorem to (8) will determine the spectrum of the fluctuations of scale \mathbf{q} in electron flux. Letting $|\Delta S_i^e(\mathbf{q}, \omega)|^2$ denote the power spectrum of the component in direction i of these fluctuations, we have, from (4) and (8),

$$\begin{aligned} & \overline{|\Delta S_i^e(\mathbf{q}, \omega)|^2} d\omega \\ &= \frac{2KT}{\pi V} \text{Real } Y'_{ii}(\mathbf{q}, \omega) d\omega \end{aligned} \tag{10}$$

To obtain the spectrum of the fluctuations in electron density, it is noted from the continuity equation,

$$\nabla \cdot \mathbf{S}^e = -\frac{\partial N^e}{\partial t}$$

where N^e denotes the electron density, that a

fluctuation

$$\Delta \mathbf{S}^e(\mathbf{q}, \omega) \exp [j(\omega t - \mathbf{q} \cdot \mathbf{r})]$$

in electron flux is accompanied by an electron density fluctuation

$$\begin{aligned} & \Delta N(\mathbf{q}, \omega) \exp [j(\omega t - \mathbf{q} \cdot \mathbf{r})] \\ &= \frac{q}{\omega} \Delta S_q^e(\mathbf{q}, \omega) \exp [j(\omega t - \mathbf{q} \cdot \mathbf{r})] \end{aligned}$$

where

$$q \Delta S_q^e(\mathbf{q}, \omega) = \mathbf{q} \cdot \Delta \mathbf{S}^e(\mathbf{q}, \omega)$$

It follows that the power spectrum of the fluctuations of scale \mathbf{q} in electron density related to the power spectrum of the fluctuations of scale \mathbf{q} in the electron flux by

$$\overline{|\Delta N(\mathbf{q}, \omega)|^2} = \frac{q^2}{\omega^2} \overline{|\Delta S_q^e(\mathbf{q}, \omega)|^2}$$

Therefore, from (10),

$$\begin{aligned} & \overline{|\Delta N(\mathbf{q}, \omega)|^2} d\omega \\ &= \frac{2KT}{\pi V \omega^2} \text{Real} (\mathbf{q} \cdot \mathbf{Y}' \cdot \mathbf{q}) d\omega \end{aligned}$$

The spectral differential scattering cross section per unit volume, equation 1, may now be rewritten by substituting for $|\Delta N(\mathbf{q}, \omega)|^2$ equation 12

$$\begin{aligned} & \sigma(\omega_0 \pm \omega) d\omega \\ &= r_e^2 \sin^2 \xi \frac{KT q^2}{\pi \omega^2} \text{Real } Y'_{zz} d\omega \end{aligned}$$

where

$$\mathbf{q} \cdot \mathbf{Y}' \cdot \mathbf{q} = q^2 Y'_{zz}$$

Here, the elements of the admittance tensor \mathbf{Y}^e , \mathbf{Y}^i , and \mathbf{Y}' are expressed in a rectangular coordinate system x, y, z , with z axis in the direction of \mathbf{q} . It proves convenient to introduce dimensionless admittance tensors \mathbf{y}^e and \mathbf{y}^i defined by

$$Y_{ii}^{e,i} \equiv \frac{\bar{N} \omega}{KT q^2} y_{ii}^{e,i}$$

where the mean electron density \bar{N}^e is equal to the mean ion density \bar{N}^i and

$$\bar{N}^e = \bar{N}^i \equiv \bar{N}$$

larly, a dimensionless matrix \mathbf{g} , related matrix in (7), is defined by

$$\frac{\mathbf{G}_{ij}}{c^2} = \frac{\bar{N}\omega}{KTq^2} \mathbf{g}_{ij} \quad (15)$$

When equations 14, 15, and 9 are substituted in equation 13, we obtain

$$\sigma(\omega) d\omega = \bar{N}r_e^2 \sin^2 \xi \operatorname{Real} \left(\frac{y'_{zz}}{\pi\omega} \right) d\omega \quad (16)$$

$$[(\mathbf{y}^i - \mathbf{g}) \cdot (\mathbf{y}^e + \mathbf{y}^i - \mathbf{g})^{-1} \cdot \mathbf{y}^e]_{zz} \quad (17)$$

tensors \mathbf{y}^i and \mathbf{y}^e are obtained by using the electron mass respectively in equation (17) Appendix B. The \mathbf{g} matrix is diagonal in y, z system with elements

$$g_{yy} = j d_0^2 q^2 \left(\frac{c^2 q^2}{\omega^2} - 1 \right) \approx j(d_0^2 q^2) \frac{c^2 q^2}{\omega^2}$$

$$\gamma_{zz} = -j d_0^2 q^2$$

bye length is defined by

$$d_0 = \left[\frac{KT}{4\pi \bar{N}e^2} \right]^{1/2} \quad (18)$$

Equation 16 is the formal solution to the problem. It is a restatement of the result of Dougherty and Farley [1960] equations (16) and (17). There are two significant differences from this earlier work, however. First, the validity of equation 1 in the presence of a magnetic field has been demonstrated in Appendix A. Second, the assumption that the admittance tensor would have to be diagonalized has been shown to be unnecessary—a conclusion that has also been reached by Dougherty and Farley in the draft copy of their paper on this topic.

Equation 16 is still formidable, since the dimensionless effective admittance function, y'_{zz} , consists of nine terms, each term a triple product of elements of $(\mathbf{y}^i - \mathbf{g})$, $(\mathbf{y}^i + \mathbf{y}^e - \mathbf{g})^{-1}$ and each element a complicated function. In order to reduce this solution to a more amenable form, Dougherty and Farley (draft copy) give a criterion for the validity of an approximation to the effective admittance tensor. We obtain a criterion which is more restrictive by the factor $(m^i/m^e)^{1/2}$. The proof of this approxi-

mation and its range of validity are given in Appendix C. Within the validity of this criterion, for the case of ionospheric backscattering only the y_{zz}^e and y_{zz}^i components of the admittance tensors are required for a good approximation of y'_{zz} , i.e.,

$$y'_{zz} = \frac{(y_{zz}^i + j d_0^2 q^2) y_{zz}^e}{y_{zz}^i + y_{zz}^e + j d_0^2 q^2} \quad (19)$$

provided that

$$\delta \leq 89.8^\circ \text{ when the ion mass is that of the oxygen atom} \quad (20)$$

$$\delta \leq 89.4^\circ \text{ when the ion mass is that of the proton}$$

where δ is the acute angle between \mathbf{q} and the direction of the external magnetic field.

According to Appendix B

$$y_{zz}^i = j + \theta^i \int_0^\infty dt \exp \left\{ -j\theta^i t - \frac{\sin^2 \delta}{(\phi^i)^2} \sin^2 \left(\frac{\phi^i t}{2} \right) - \frac{t^2 \cos^2 \delta}{4} \right\} \quad (21)$$

$$y_{zz}^e = j + \theta^e \int_0^\infty dt \exp \left\{ -j\theta^e t - \frac{\sin^2 \delta}{(\phi^e)^2} \sin^2 \left(\frac{\phi^e t}{2} \right) - \frac{t^2 \cos^2 \delta}{4} \right\} \quad (22)$$

$$\theta \equiv \frac{\omega}{q} \left(\frac{m}{2KT} \right)^{1/2} \quad (23)$$

$$\phi = \frac{\Omega}{q} \left(\frac{m}{2KT} \right)^{1/2} = \frac{1}{qR} \quad (24)$$

$$\Omega = \frac{eB_0}{mc} = \text{angular gyrofrequency} \quad (25)$$

θ represents a dimensionless ratio of Doppler shift to operating frequency. ϕ is the dimensionless gyrofrequency, and R will be referred to as the gyroradius. ϕ and θ take superscripts i or e corresponding to positive ion or electron mass respectively.

The spectrum equation 16 reduces, within the validity of criterion (20), to

$$\sigma(\omega_0 \pm \omega) d\omega = \frac{\bar{N}r_e^2 \sin^2 \xi}{\sqrt{\pi}} \cdot \left\{ \operatorname{Real} \left[\frac{(y_{zz}^i + j d_0^2 q^2) y_{zz}^e}{y_{zz}^i + y_{zz}^e + j d_0^2 q^2} \right] \right\} \frac{d\omega}{\sqrt{\pi\omega}} \quad (26)$$

In the next section, the behavior of (26) is studied for various values of the parameters, and plots are presented for the special case of backscattering.

3. ANALYTICAL EXPRESSIONS AND CORRESPONDING GRAPHICAL REPRESENTATIONS OF THE BACKSCATTER SPECTRUM

In the analysis which follows, we particularize to the case of backscattering by letting $\xi = \pi/2$, $\epsilon = \pi$. Then

$$\begin{aligned} |q| &= 4\pi/\lambda = 2\omega_0/c \\ \phi^{e,i} &= (\lambda/4\pi)(1/R^{e,i}) \\ \theta^{e,i} &= \frac{\omega}{2\omega_0} \left(\frac{m^{e,i}c^2}{2KT} \right)^{1/2} \\ \cos \delta &= \frac{\mathbf{k}_{\text{inc}} \cdot \mathbf{B}_0}{|\mathbf{k}_{\text{inc}}| |\mathbf{B}_0|} \end{aligned}$$

Let y_{zz}^e and y_{zz}^i be expressed by

$$\begin{aligned} y_{zz}^e &= R_e(\theta^e, \phi^e, \delta) + jI_e(\theta^e, \phi^e, \delta) \\ &\equiv R_e + jI_e \end{aligned} \quad (27)$$

$$\begin{aligned} y_{zz}^i &= R_i(\theta^i, \phi^i, \delta) + jI_i(\theta^i, \phi^i, \delta) \\ &\equiv R_i + jI_i \end{aligned} \quad (28)$$

R_e , R_i , I_e , and I_i are the real and imaginary parts of y_{zz}^e and y_{zz}^i and are not to be confused with R^e and R^i , which denote the electron and positive ion gyroradii respectively. Combining (26), (27), and (28), we obtain

$$\chi(\omega) d\omega = \frac{R_e R_i (R_e + R_i) + R_e (I_e + d_0^2 q^2)^2 + R_i I_e^2}{(R_e + R_i)^2 + (I_e + I_i + d_0^2 q^2)^2} \frac{d\theta^i}{\sqrt{\pi} \theta^i}$$

where

$$\chi(\omega) = \frac{\sqrt{\pi}}{N r_e^2} \sigma(\omega_0 \pm \omega)$$

The right-hand side of equation 29a is a function of θ^i , since

$$\theta^e = \theta^i \left(\frac{m^i}{m^e} \right)^{1/2}$$

It should further be noted that equation 29a reduces to

$$\chi(\omega) d\omega = R_e \frac{d\theta^i}{\sqrt{\pi} \theta^i} = R_e \frac{d\theta^e}{\sqrt{\pi} \theta^e} \quad (29b)$$

when $d_0^2 q^2 \gg I_e + I_i$ and

$$d_0^4 q^4 \gg (R_e + R_i)^2, \quad |R_i(R_i + R_e)|, \quad \left| \frac{I_e}{R_e} \right|$$

These conditions are satisfied for $d_0 q \geq 10 \lambda \leq d_0$.

The gyroradius, R^i , of the ions is equal to the gyroradius, R^e , of the electrons times $(m^i/m^e)^{1/2}$ so that, when the operating wavelength, λ , is of the order of $4\pi R^e$, λ is necessarily much smaller than $4\pi R^i$, and, conversely, if λ is of the order of $4\pi R^i$, λ is necessarily much greater than $4\pi R^e$. For a given set of parameters, d_0 , R^e , and R^i , approximate expressions can be obtained for the spectrum in the five regions indicated below.

- Case 1. $\lambda/4\pi \ll R^e \ll R^i$.
- Case 2. $\lambda/4\pi \sim R^e \ll R^i$.
- Case 3. $R^e \ll \lambda/4\pi \ll R^i$.
- Case 4. $R^e \ll \lambda/4\pi \sim R^i$.
- Case 5. $R^e \ll R^i \ll \lambda/4\pi$.

For each of these five cases, the ratio of the Debye length, d_0 , to the wavelength, λ , is a variable.

Case 1. $\lambda/4\pi \ll R^e \ll R^i$; i.e., ϕ^e and ϕ^i are much smaller than unity, equivalent to a vanishingly small \mathbf{B}_0 field, the quantities

$$\frac{\sin^2 \delta \sin^2(\phi^i t/2)}{(\phi^i)^2}, \quad \frac{\sin^2 \delta \sin^2(\phi^e t/2)}{(\phi^e)^2}$$

reduce to $(t^2/4) \sin^2 \delta$ in the integrands of equations 21 and 22. Thus

$$y_{zz}^e = j + \theta^e \int_0^\infty \exp[-j\theta^e t - \tfrac{1}{4}t^2] dt$$

$$y_{zz}^i = j + \theta^i \int_0^\infty \exp[-j\theta^i t - \tfrac{1}{4}t^2] dt$$

Integrating (30) and (31),

$$\begin{aligned} y_{zz}^e &= \sqrt{\pi} \theta^e e^{-(\theta^e)^2} \\ &\quad + j \left\{ 1 - 2\theta^e e^{-(\theta^e)^2} \int_0^{\theta^e} e^{p^2} dp \right\} \\ y_{zz}^i &= \sqrt{\pi} \theta^i e^{-(\theta^i)^2} \\ &\quad + j \left\{ 1 - 2\theta^i e^{-(\theta^i)^2} \int_0^{\theta^i} e^{p^2} dp \right\} \end{aligned}$$

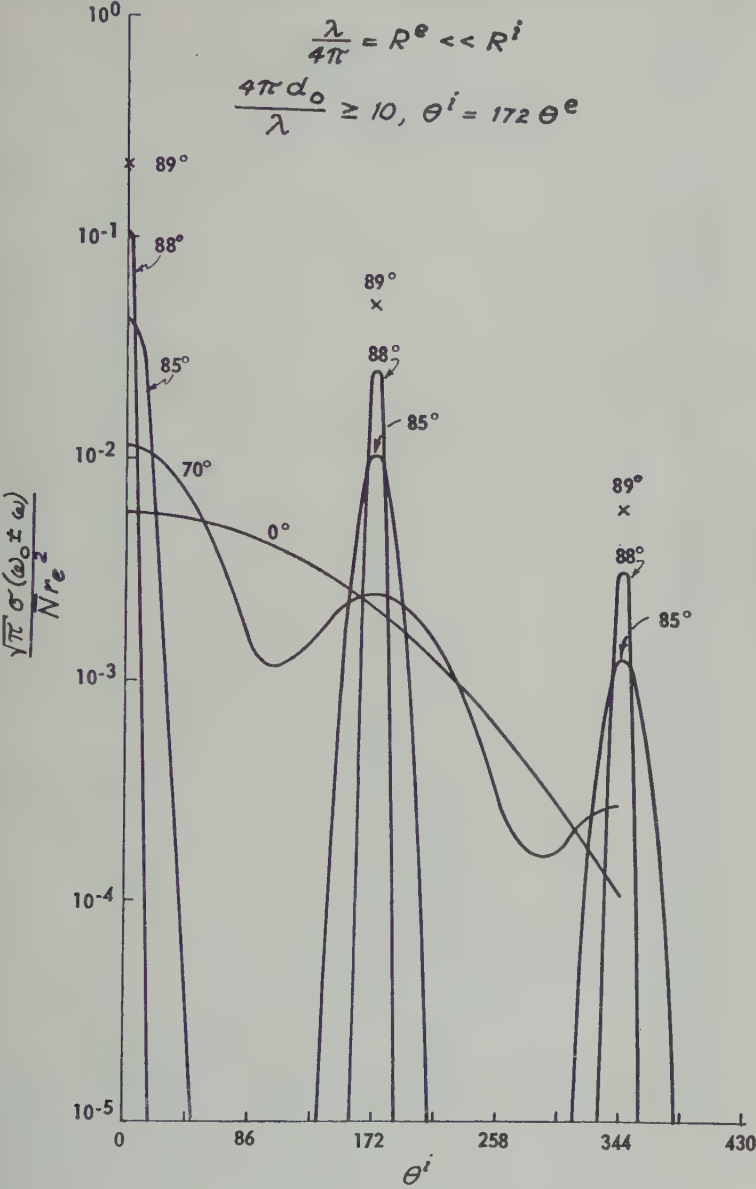


Fig. 1. Spectrum of backscattered echoes.

numerical calculations we will use

$$\int_0^z e^{p^2} dp \approx 2z^2 \quad \text{for } z \leq 0.2$$

$$\approx 1 + \frac{1}{2z^2} \quad \text{for } z > 2 \quad (34)$$

Comparing (27) and (28) with (32) and (33)

$$R_e = \sqrt{\pi \theta^e} e^{-(\theta^e)^2} \quad (35a)$$

$$I_e = 1 - 2\theta^e e^{-(\theta^e)^2} \int_0^{\theta^e} e^{p^2} dp \quad (35b)$$

$$R_i = \sqrt{\pi \theta^i} e^{-(\theta^i)^2} \quad (36a)$$

$$I_i = 1 - 2\theta^i e^{-(\theta^i)^2} \int_0^{\theta^i} e^{p^2} dp \quad (36b)$$

When equations 35a, 35b, 36a, and 36b are substituted into equation 29a, the resulting

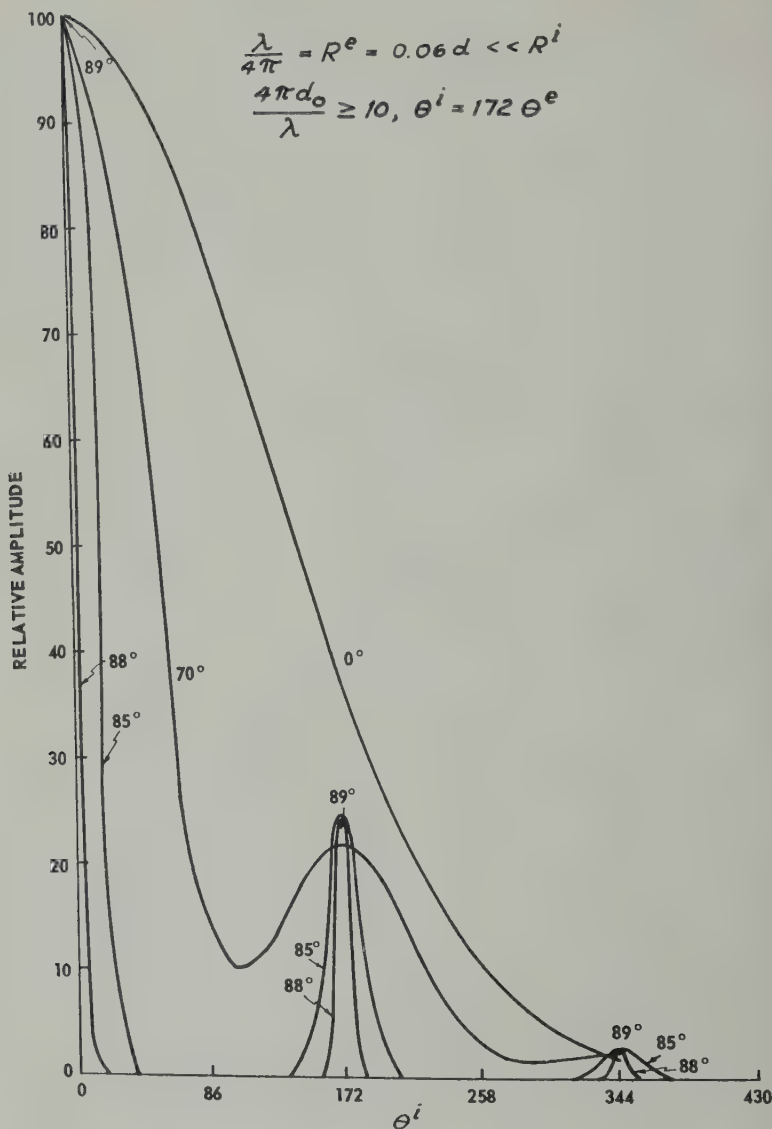


Fig. 2. Spectrum of backscattered echoes.

spectrum function is identical to that obtained by *Dougherty and Farley* [1960] for the magnetic field-free case. Since the result is independent of the angle δ , we refer to this case as the isotropic case. It should further be noted that, when δ in equations 21 and 22 is equal to zero, that is, when the direction of propagation is parallel to the magnetic field, the backscatter spectrum function is identical to the spectrum of the isotropic case for all values of ϕ^e and ϕ^i . The

graphical results for case 1 are illustrated in curves for cases 2 to 4 by the $\delta = 0$ values.

For $\lambda \leq d_0$, the introduction of R_e , equation 35a, into the simplified expression 29b yields

$$\chi(\omega) d\omega = \exp \{ -(\theta^e)^2 \} d\theta^e$$

which is the same Gaussian function that obtained for the scatter spectrum arising from noninteracting electrons in thermal equilibrium.

Case 2. $\lambda/4\pi \sim R^e \ll R^i$; i.e., $\phi^e \sim 1, \phi^i \sim 1$.

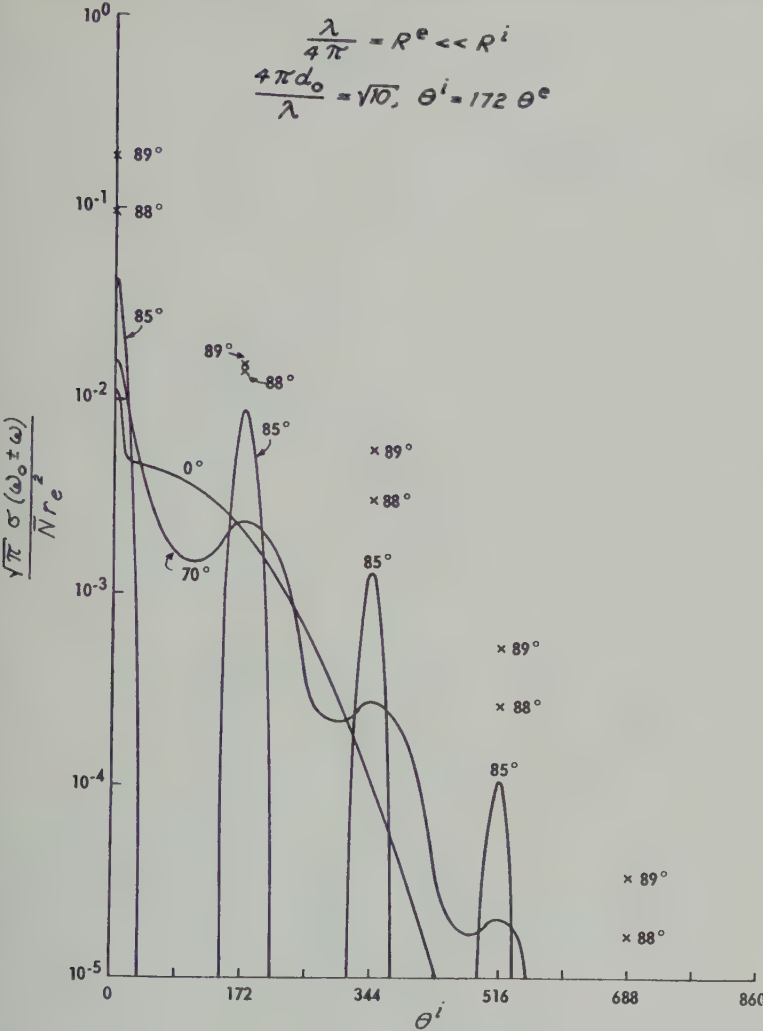


Fig. 3. Spectrum of backscattered echoes.

θ^i is still much smaller than unity, equations 36a and 36b will still represent R_e and I_e . However, expressions for R_e and I_e valid for ϕ^e of order of unity must be derived. To do this identity

$$\sum_{n=-\infty}^{\infty} j^n J_n(-jl) e^{jn\phi^e t} = \sum_{n=-\infty}^{\infty} j^{-n} J_n(jl) e^{jn\phi^e t} = \sum_{n=-\infty}^{\infty} I_n(l) e^{jn\phi^e t} \quad (38)$$

$j^{-n} J_n(jl) = I_n(l)$ is the Bessel function of imaginary argument. (See Dwight [1955], equation 3.3, for the series expression of $I_n(l)$.)

Therefore

$$\exp \left[-\frac{\sin^2 \delta \sin^2 (\phi^e t/2)}{(\phi^e)^2} \right] = e^{-l} \sum_{n=-\infty}^{\infty} I_n(l) e^{jn\phi^e t} \quad (39)$$

where

$$l = \frac{\sin^2 \delta}{2(\phi^e)^2}$$

Substituting equation 39 into equation 22, and interchanging the summation and the integration signs, we obtain

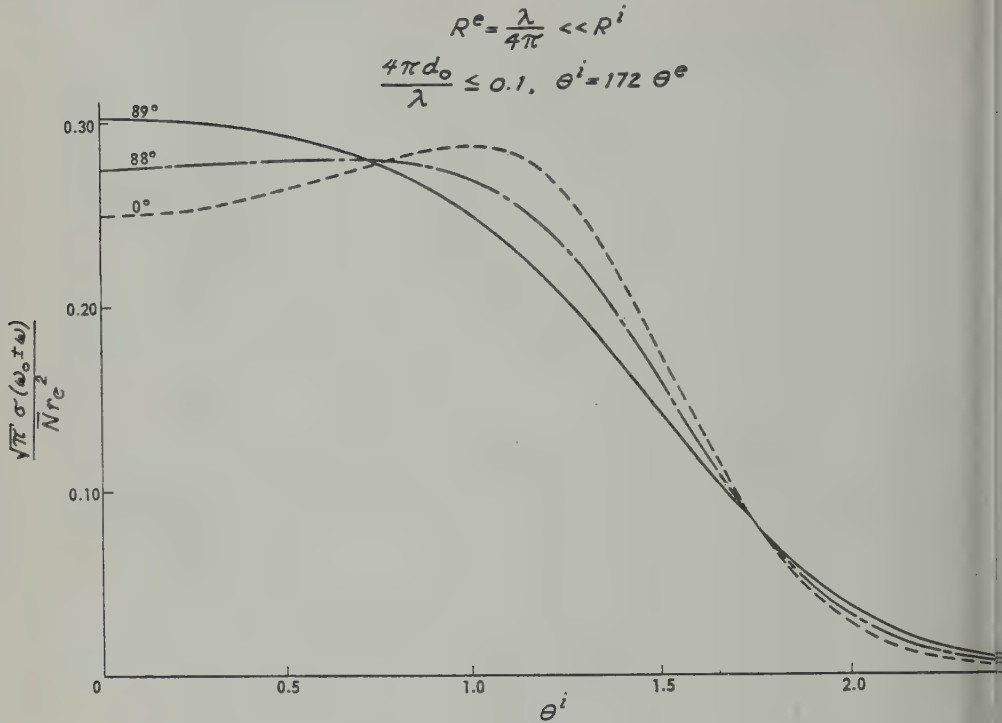


Fig. 4. Spectrum of backscattered echoes.

$$y_{zz}^e = j + \theta^e e^{-1} \sum_{n=-\infty}^{\infty} I_n(l)$$

$$\cdot \int_0^{\infty} \exp \left(-j\theta^e t + jn\phi^e t - \frac{t^2 \cos^2 \delta}{4} \right) dt \quad (40)$$

The integration of the real and imaginary parts of y_{zz}^e yields

$$R_e = \frac{\sqrt{\pi} \theta^e}{\cos \delta} \exp \left(-\frac{\sin^2 \delta}{2(\phi^e)^2} \right)$$

$$\cdot \sum_{n=-\infty}^{\infty} I_n \left[\frac{\sin^2 \delta}{2(\phi^e)^2} \right] \exp \left[-\left(\frac{\theta^e - n\phi^e}{\cos \delta} \right)^2 \right], \quad (41a)$$

$$I_e = 1 - \frac{2\theta^e \exp \left[-\frac{\sin^2 \delta}{2(\theta^e)^2} \right]}{\cos \delta}$$

$$\cdot \sum_{n=-\infty}^{\infty} I_n \left[\frac{\sin^2 \delta}{2(\phi^e)^2} \right] \exp \left[-\left(\frac{\theta^e - n\phi^e}{\cos \delta} \right)^2 \right]$$

$$\cdot \int_0^{(\theta^e - n\phi^e)/(\cos \delta)} e^{p^2} dp \quad (41b)$$

The substitution of equations 36a, 36b, and 41b into equation 29a results in the spectrum function. For $\lambda \leq d_0$, it is sufficient to introduce equation 41a into 29b; the resulting function contains the results reported by Laaspere [19].

The results of our calculations of equation 29a for case 2, and for various values of δ , are illustrated in Figures 1 to 6. In each case the spectrum function $\chi(\omega)$ is plotted as a function of θ^i for parametric values of δ . The spectrum function is not shown for negative θ^i as it is symmetric about $\theta^i = 0$. For $\lambda \gg d_0$ (Fig. 4) most of the scattered energy is contained in the range $0 \leq \theta^i < 3$. The spectrum shows no peaks at $\theta^i = 172, 344$, etc., but these are reduced in magnitude by a factor of at least 10 from those of Figure 2.

The peaks and troughs in Figures 1, 2, 4, and 6 occur approximately at multiples of the electron gyrofrequency. When the wavelength is smaller than the Debye length, the peaks occur almost exactly at multiples of the gyrofrequency. As the wavelength becomes greater than the Debye length, the peaks are shifted away from the positions of the harmonics of the

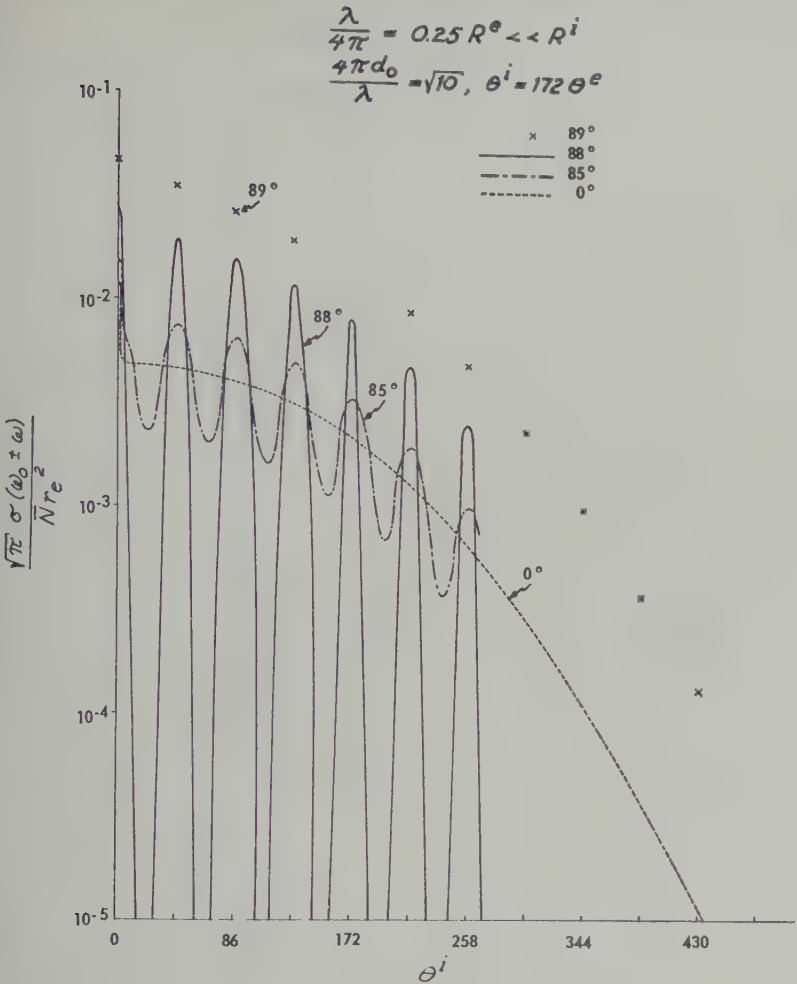


Fig. 5. Spectrum of backscattered echoes.

ey. To illustrate this point, the detail of
k near $\theta^i = 86$ and $\delta = 85^\circ$ in the
n of Figure 5 has been plotted on an
d scale in Figure 7. Additional curves
spectra for $4\pi d_0/\lambda = 15$ and 1 are also
in the neighborhood of $\theta^i = 86$ in the
figure 7. For $4\pi d_0/\lambda = 15$ ($\lambda \lesssim d_0$)
n 41a would lead one to expect peaks at
 ϕ^e where $n = \mp 1, \mp 2, \mp 3$, etc.
ks near $\theta^i = 86$ (Fig. 7) correspond to
since $\phi^e = 0.25$. The curve for $\lambda \lesssim d_0$
k at $\theta^i = 86$. As the wavelength increases
comes larger than the Debye length,
, the peak shifts to the right, as illus-
y the curves corresponding to $4\pi d_0/\lambda =$

$\sqrt{10}$ and $4\pi d_0/\lambda = 1$. The Coulomb interaction
effects, important for $\lambda > d_0$, cause the peaks of
the spectrum to shift from the values that
would be calculated on a single-particle basis.
Figure 8 is a plot of the relative values of the
peaks of the spectra for parametric values of ϕ^e
and fixed values of δ and λ . For comparison, the
plot of the Gaussian spectrum of a field-free
ionized gas for $\lambda \lesssim d_0$ is given. The other spectra
have been renormalized to coincide with the
Gaussian spectrum at $\theta^i = 0$, and so the plot
clearly indicates the deviations from the Gaussian
spectrum produced by the magnetic field. For
 $\phi^e \leq 0.25$ the agreement with the Gaussian
curve is quite good. However, it can be shown
analytically, using the asymptotic form of equa-

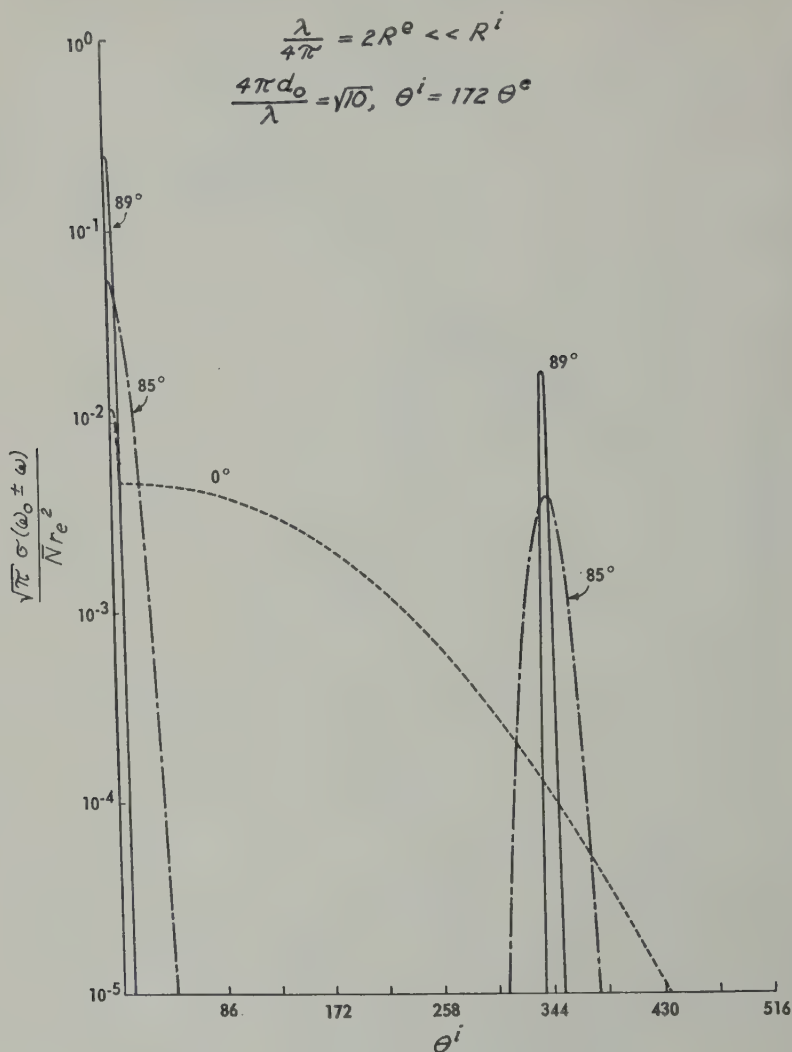


Fig. 6. Spectrum of backscattered echoes.

tion 41a, that, as λ becomes increasingly greater than R^e , i.e., $\phi^e > 0.25$, the spectrum envelope increasingly deviates from the Gaussian spectrum.

Case 3. $R^e \ll \lambda/4\pi \ll R^i$; i.e., $\phi^e \gg 1$, $\phi^i \ll 1$. As in cases 1 and 2, ϕ^i is still much less than unity, and so (36a) and (36b) are still applicable. To obtain R_e and I_e for $\phi^e \gg 1$, the integration of equation 22 can be simplified by noting that

$$\begin{aligned} & \exp \left[-\frac{\sin^2 \delta}{(\phi^e)^2} \sin^2 \left(\frac{\phi^e t}{2} \right) \right] \\ &= \exp \left[-\frac{\sin^2 \delta}{2(\phi^e)^2} \right] \exp \left[\frac{\sin^2 \delta \cos \phi^e t}{2(\phi^e)^2} \right] \end{aligned}$$

$$\approx \left(1 + \frac{\sin^2 \delta \cos \phi^e t}{2(\phi^e)^2} \right) \exp \left[-\frac{\sin^2 \delta}{2(\phi^e)^2} \right]$$

The substitution of equation 42 into equation 22 leads to

$$\begin{aligned} y_{zz}^e &= j + \theta^e \exp \left[-\frac{\sin^2 \delta}{2(\phi^e)^2} \right] \\ &\cdot \left\{ \exp \left[-\left(\frac{\theta^e}{\cos \delta} \right)^2 \right] \int_0^\infty \exp \left[j \left(\theta^e t - \frac{t^2 \cos^2 \delta}{4} \right) \right] \right. \\ &\quad \left. - \frac{\theta^e}{\cos \delta} \right]^2 dt + \frac{\sin^2 \delta}{2(\phi^e)^2} \int_0^\infty \cos(\phi^e t) \cdot \exp \left[-j\theta^e t - \frac{t^2 \cos^2 \delta}{4} \right] dt \right\} \end{aligned}$$

$$\frac{\lambda}{4\pi} = 0.25 R^e \ll R^i$$

$$\theta^i = 172^\circ, \delta = 85^\circ$$

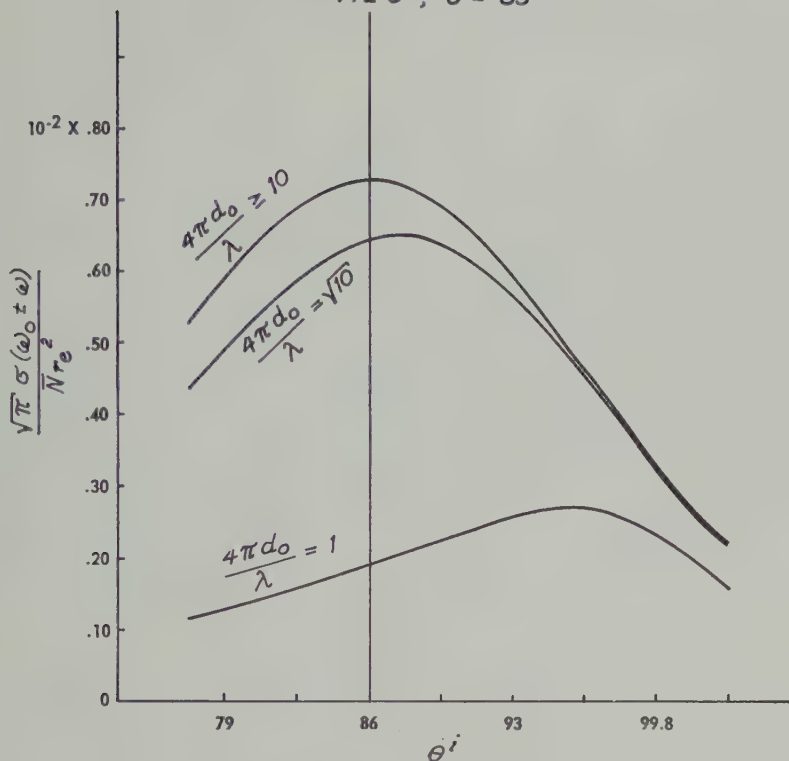


Fig. 7. Spectrum of backscattered echoes.

when integrated results in

$$= \frac{\sqrt{\pi} \theta^e}{\cos \delta} \left\{ \exp \left[-\left(\frac{\theta^e}{\cos \delta} \right)^2 \right] + \frac{\sin^2 \delta}{4(\phi^e)^2} \right. \\ \left. \left[\exp \left(-\left(\frac{\theta^e - \phi^e}{\cos \delta} \right)^2 \right) \right. \right. \right. \\ \left. \left. \left. + \exp \left(-\left(\frac{\theta^e + \phi^e}{\cos \delta} \right)^2 \right) \right] \right\} \exp \left(-\frac{\sin^2 \delta}{2(\phi^e)^2} \right) \quad (43a)$$

$$1 - \frac{2\theta^e}{\cos \delta} \left\{ \exp \left[-\left(\frac{\theta^e}{\cos \delta} \right)^2 \right] \right. \\ \left. \int_0^{\theta^e / \cos \delta} e^{p^2} dp + \frac{\sin^2 \delta}{4(\phi^e)^2} \exp \left(-\frac{\theta^e - \phi^e}{\cos \delta} \right) \right. \\ \left. \int_0^{(\theta^e - \phi^e) / \cos \delta} e^{p^2} dp + \frac{\sin^2 \delta}{4(\phi^e)^2} \right. \\ \left. \exp \left[-\left(\frac{\theta^e + \phi^e}{\cos \delta} \right)^2 \right] \int_0^{(\theta^e + \phi^e) / \cos \delta} e^{p^2} dp \right\}$$

$$\cdot \exp \left(-\frac{\sin^2 \delta}{2(\phi^e)^2} \right) \quad (43b)$$

For case 3, consider the behavior of the spectrum function in the two regions $0 \leq \theta^e \leq 2$ and $\theta^e > 2$, i.e., $0 \leq \theta^i \leq 344$ and $\theta^i > 344$.

(a) $\theta^e \leq 2$; $\phi^e \gg 1$, $\phi^i \ll 1$. Equations 43a and 43b reduce to

$$R_e = \sqrt{\pi} \frac{\theta^e}{\cos \delta} \exp \left[-\left(\frac{\theta^e}{\cos \delta} \right)^2 \right] \quad (44a)$$

$$I_e = 1 - \frac{2\theta^e}{\cos \delta} \\ \cdot \exp \left[-\left(\frac{\theta^e}{\cos \delta} \right)^2 \right] \int_0^{\theta^e / (\cos \delta)} e^{p^2} dp \quad (44b)$$

The substitution of equations 44a, 44b, 36a, and 36b into 29a will give the explicit spectrum function for $\theta^i \leq 344$. Figures 9, 10, and 11 are

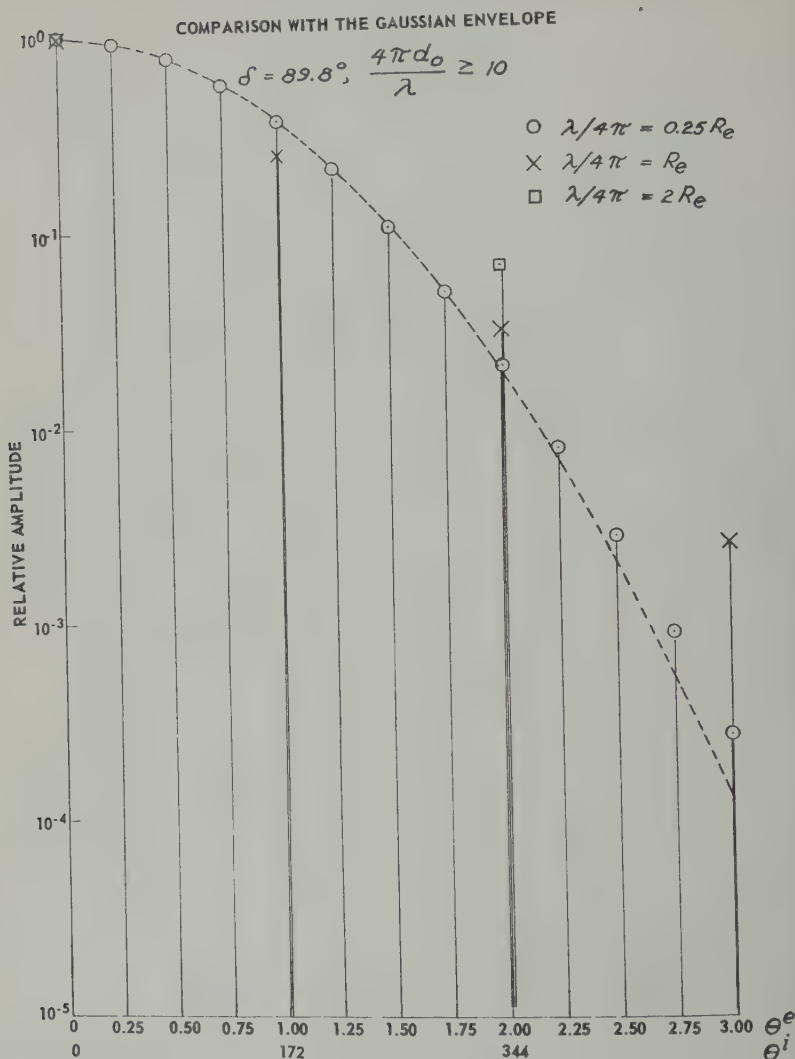


Fig. 8. Spectrum of backscattered echoes.

the results of the calculations for case 3, with $\lambda \gg d_0$, $\lambda \approx 4d_0$, and $\lambda \leq d_0$, respectively. In Figures 12 and 13 the results are plotted on logarithmic scales in order to emphasize the variation of the spectrum function with increasing δ . The hump that is present in the $\delta = 0$ curve of Figure 12 ($\lambda \approx 4d_0$) is due to the weakened, but still present, Coulomb interaction; the hump vanishes for $\lambda \leq d_0$ as shown in Figure 13.

(b) $2 < \theta^e \ll \phi^e$, i.e., $344 < \theta^e \ll 172^\circ$; $\phi^e \gg 1$, $\phi^i \ll 1$. R_i and I_i (equations 36a and

36b) are of the order of zero, so that the spectrum function in this region has the form

$$\chi d\omega]_{\theta^e > 2} \approx \left\{ \frac{R_e d_0^4 q^4}{R_e^2 + [I_e + d_0^2 q^2]^2} \right\} \frac{d\omega}{\sqrt{\dots}}$$

Equations 44a and 44b are still valid for R_e and I_e . Moreover, for equation 44b, with the aid of (34), reduce

$$I_e \approx \frac{-\cos^2 \delta}{2(\theta^e)^2}$$

Substituting equations 44a and 46 into

$$\chi(\theta^e > 2) \approx \frac{d_0^4 q^4 \theta^e \exp \left[-\left(\frac{\theta^e}{\cos \delta} \right)^2 \right] d\omega}{\omega \sqrt{\pi} \cos \delta \left\{ \pi \left(\frac{\theta^e}{\cos \delta} \right)^2 \exp \left[-2 \left(\frac{\theta^e}{\cos \delta} \right)^2 \right] + d_0^4 q^4 \left[1 - \frac{\cos^2 \delta}{2(\theta^e)^2 d_0^2 q^2} \right]^2 \right\}} \quad (47)$$

the $d\theta^i/\theta^i = d\omega/\omega$.

There is a peak in the spectrum when

$$2(\theta^e)^2 d_0^2 q^2 = \cos^2 \delta \quad (48)$$

$d_0^2 q^2 \ll 1$, since $\theta^e > 2$.

With the definitions of θ^e , q , and d_0 , (48)

can be rewritten as

$$\omega = \omega_{pe} \cos \delta$$

where $\omega_{pe} = 2\pi$ times the electron plasma frequency, f_{pe} . Therefore, the resonance occurs at the carrier frequency plus the effective plasma frequency $f_{pe} \cos \delta$.

The spectrum amplitude at the resonant frequency is

$$\chi(\omega)_{\text{resonance}} \approx \sqrt{\frac{2}{\pi}} \frac{(d_0 q)^5}{\pi}$$

$$\cdot \exp \left\{ + \frac{1}{2 d_0^2 q^2} \right\} \frac{d\omega}{\omega_{pe} \cos \delta} \quad (49)$$

and the peak at resonance far exceeds the value of the spectrum at $\theta^e = \theta^i = 0$, since

$$\frac{\chi_{\text{resonance}} d\omega}{\chi(0) d\omega} = \frac{4 d_0^6 q^6 e^{+1/(2d_0^2 q^2)}}{1 + \cos \delta \sqrt{m^i/m^e}} \quad (50)$$

However, this peak is extremely narrow. Letting $\omega = \omega_{pe} \cos \delta + \delta\omega$ in (47) and solving for the value of $\delta\omega$ for which χ has half the value of the maximum in (49), we obtain for the total resonance width

$$2 \delta\omega = \sqrt{\frac{\pi}{2}} \frac{\omega_{pe} \cos \delta}{d_0^3 q^3} e^{-1/(2d_0^2 q^2)} \quad (51)$$

Note that the width is reduced as δ approaches

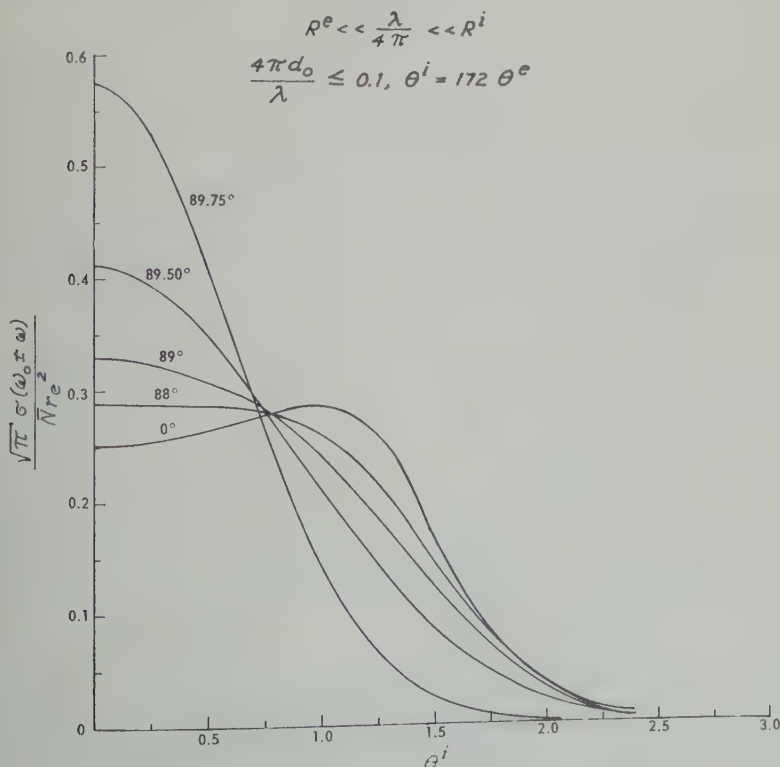


Fig. 9. Spectrum of backscattered echoes.

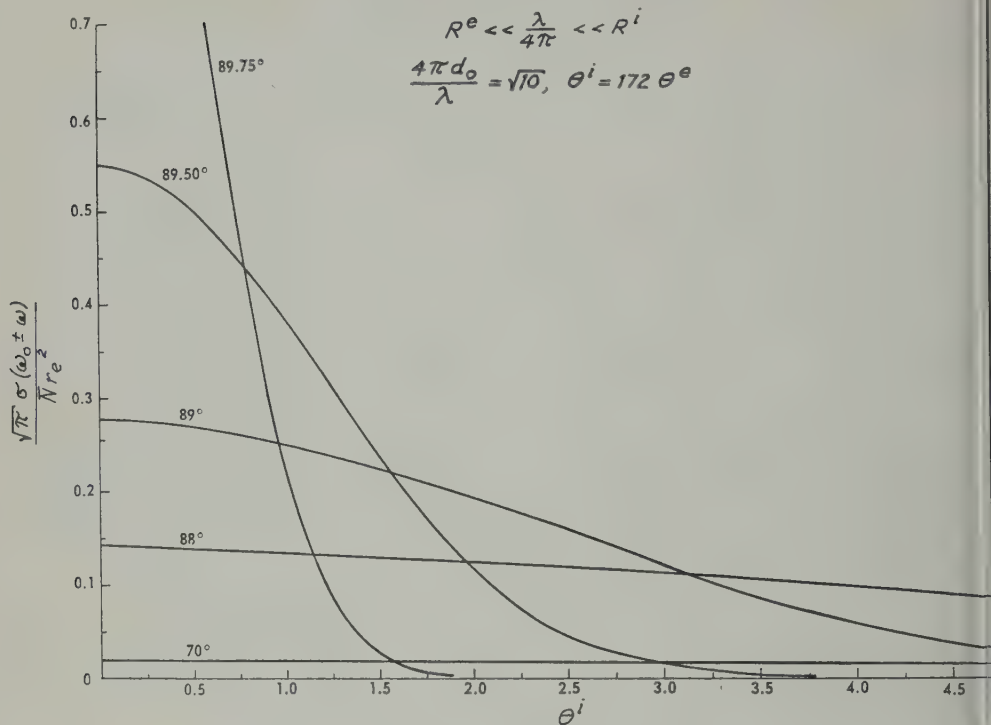


Fig. 10. Spectrum of backscattered echoes.

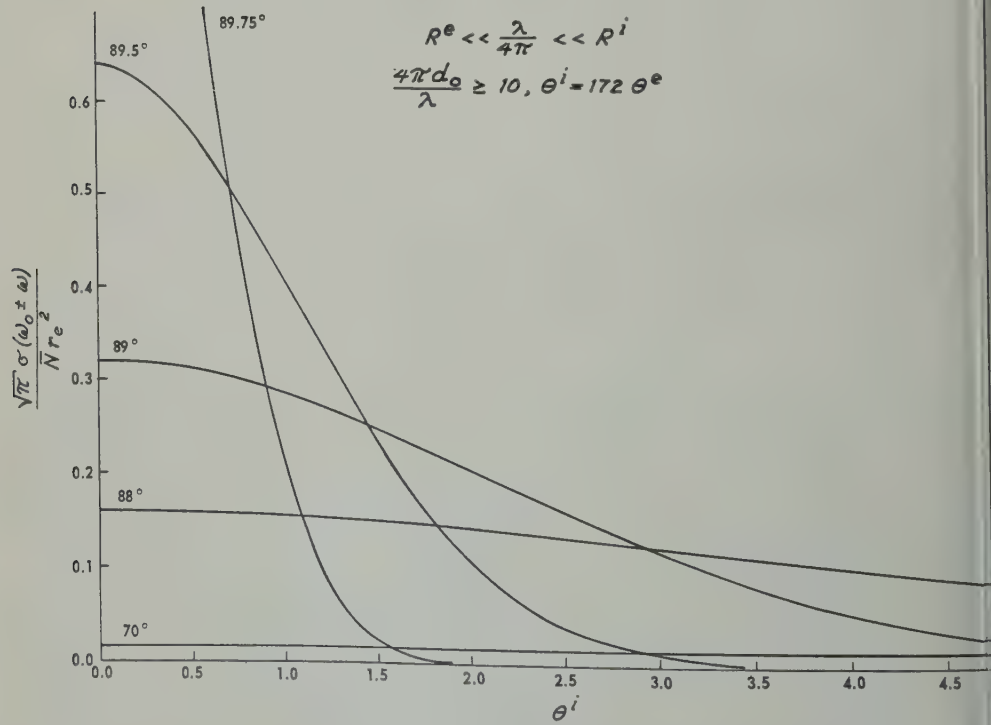


Fig. 11. Spectrum of backscattered echoes.

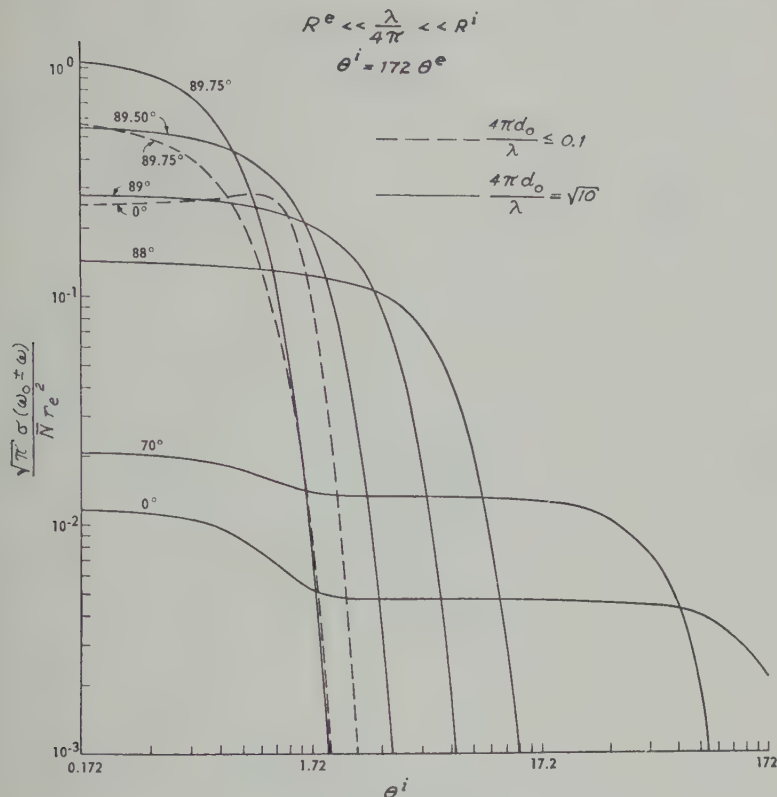


Fig. 12. Spectrum of backscattered echoes.

2. The area under this peak is approximately $\frac{1}{2}\pi$, independent of the value of δ , and is small compared with the rest of the spectrum, since $(d_0 q)^2 \ll 1$. Additional peaks can be found in the spectrum for certain values of $\theta = n\phi^e$, but these peaks are extremely narrow and small.

Resonance scattering very near the effective plasma frequency will occur whenever $\phi^e + I_e + I_i = 0$.

Case 4. $R^e \ll \lambda/4\pi \sim R^i$; i.e., $\phi^e \gg 1$, $\phi^i \sim 1$. Equations 44a and 44b are valid for R_e and I_e , and equations 41a and 41b are appropriate for I_i when ϕ^e and θ^e are replaced by ϕ^i and θ^i , respectively. The substitution of R_e , I_e , and I_i into (29a) leads to the spectrum function for this case.

The spectra are presented in Figure 14. This is the first for which protons have been used as the positive ions in the calculation. A geospheric application, where the operating frequency may be of the order of 50 Mc/s, and the conditions of case 4 will be satisfied at

exospheric levels where protons are expected to be the predominant ion species. Dr. Kenneth Bowles of the National Bureau of Standards has chosen this frequency for his forthcoming experiments to be performed near the magnetic equator.

For wavelengths greater than the Debye length, the spectra of Figure 14 illustrate the now-familiar peaks and troughs, but in the present case the peaks occur near the harmonics of the ion gyrofrequency (shown by vertical lines in Fig. 14). The displacements of the peaks from the exact harmonics of the gyrofrequency are not dependent on the mass of the ions but do depend on the Coulomb interaction.

When the wavelength is less than the Debye length, the spectrum function reduces to (29b). Substitution of (44a) into (29b) yields spectra identical to the results of case 3 with $\lambda \leq d_0$. These results have been illustrated in Figure 11.

Case 5. $R^e \ll R^i \ll \lambda/4\pi$; i.e., $\phi^e \gg 1$, $\phi^i \gg 1$. The analytic expression for this case is obtained using equations 44a and 44b for R_e and I_e .

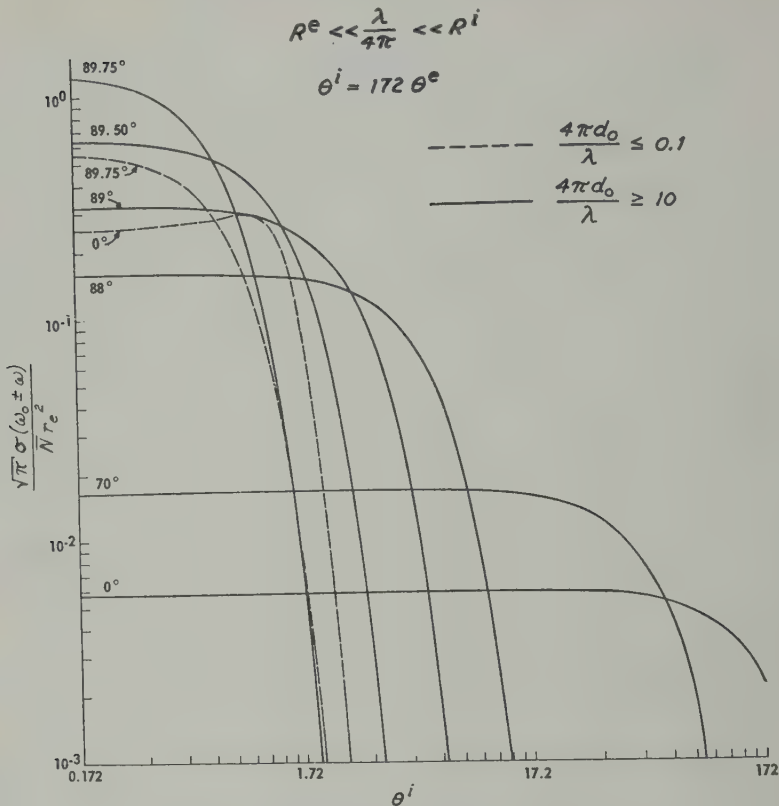


Fig. 13. Spectrum of backscattered echoes.

R_i and I_i are expressed by equations 44a and 44b with θ^e replaced by θ^i . The substitution of these expressions into (29a) provides the spectrum function for all values of d_0q . When $d_0q \geq 10$, i.e., $\lambda \leq d_0$, the simplified expression 29b can be used. Numerical computations for $\lambda \gg d_0$ are presented in Figure 15. For $\lambda \leq d_0$, the results are identical to the results of case 3 with $\lambda \leq d_0$. The values of the spectra at $\theta^i = 0$ as a function of the angle of off-perpendicularity, $\pi/2 - \delta$, in order to emphasize the aspect sensitivity for this case.

4. EVALUATION OF THE DIFFERENTIAL SCATTERING CROSS SECTION PER UNIT VOLUME

According to equation 3,

$$\sigma_T = r_e^2 \frac{V}{2} \sin^2 \xi \int_{-\infty}^{\infty} |\Delta N(q, \omega)|^2 d\omega$$

It follows from equation A17, Appendix A, that

$$\overline{\langle \Delta N_q(t) \Delta N_q^*(t + \tau) \rangle} = \frac{1}{2} \int_{-\infty}^{\infty} \overline{|\Delta N_q(\omega)|^2} d\omega$$

Putting $\tau = 0$,

$$\overline{\langle |\Delta N_q(t)|^2 \rangle} = \frac{1}{2} \int_{-\infty}^{\infty} \overline{|\Delta N_q(\omega)|^2} d\omega$$

But $\overline{|\Delta N_q(t)|^2}$ is independent of time in statistical equilibrium ensemble, and so

$$\overline{\langle |\Delta N_q(t)|^2 \rangle} = \overline{|\Delta N_q|^2}$$

where

$$\Delta N_q = \frac{1}{V} \int_V \Delta N(\mathbf{r}) \exp(j\mathbf{q} \cdot \mathbf{r}) dV$$

and $\Delta N(\mathbf{r})$ denotes the density fluctuation corresponding to a particular microscopic state of the system. From (54)

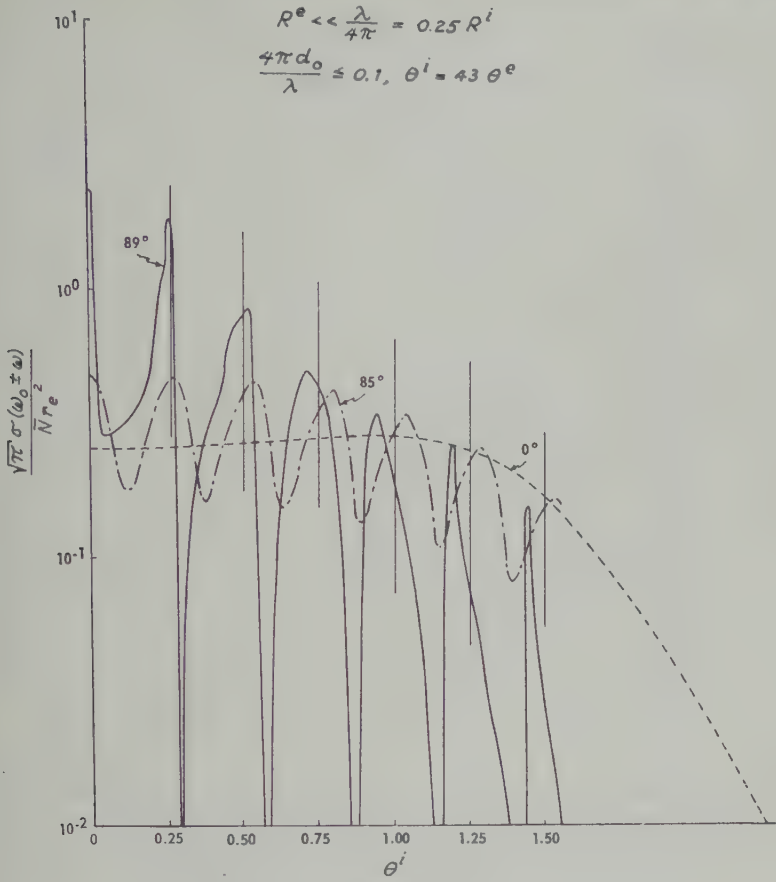


Fig. 14. Spectrum of backscattered echoes.

$$\begin{aligned} \bar{N} r_e^2 &= \frac{1}{V^2} \int_V \int_V \overline{\Delta N(\mathbf{r}) \Delta N(\mathbf{r}')} \\ &\quad \cdot \exp [j \mathbf{q} \cdot (\mathbf{r} - \mathbf{r}')] dV dV' \end{aligned}$$

Combining this relation with (52) and (53), we obtain

$$\begin{aligned} \bar{N} r_e^2 \sin^2 \xi &\int_V \int_V \overline{\Delta N(\mathbf{r}) \Delta N(\mathbf{r}')} \\ &\quad \cdot \exp [j \mathbf{q} \cdot (\mathbf{r} - \mathbf{r}')] dV dV' \end{aligned} \quad (55)$$

where the bar denotes an average over a statistical equilibrium ensemble. This result is identical to equation 1 of a paper by Renau [1960], where $\overline{(\Delta N(\mathbf{r}))^2}$ was obtained, in his Appendix A, for the case of an ionized gas in thermal equilibrium in the absence of external fields. For the case of thermal equilibrium in the presence of a static magnetic

field, it can be shown that the generalized momentum coordinates appearing in the Liouville density function are such that integration over the momentum coordinates yields the same density in configuration space used in Renau's Appendix A. Therefore, those published results are independent of the application of a static magnetic field and are applicable to the present problem. Renau (see also Salpeter [1960a and b]; Fejer [1960]; Dougherty and Farley [1960]) found that the total scattering cross section for backscattering was

$$\sigma_T = \bar{N} r_e^2 \frac{Z + \frac{d_0^2 q^2}{2}}{1 + Z + \frac{d_0^2 q^2}{2}} \quad (56)$$

where Z is the net charge of the ions. Dougherty and Farley (draft copy), in a manner originally employed in their published paper [1960], have shown that the integration of equation 16 using the approximate expression for

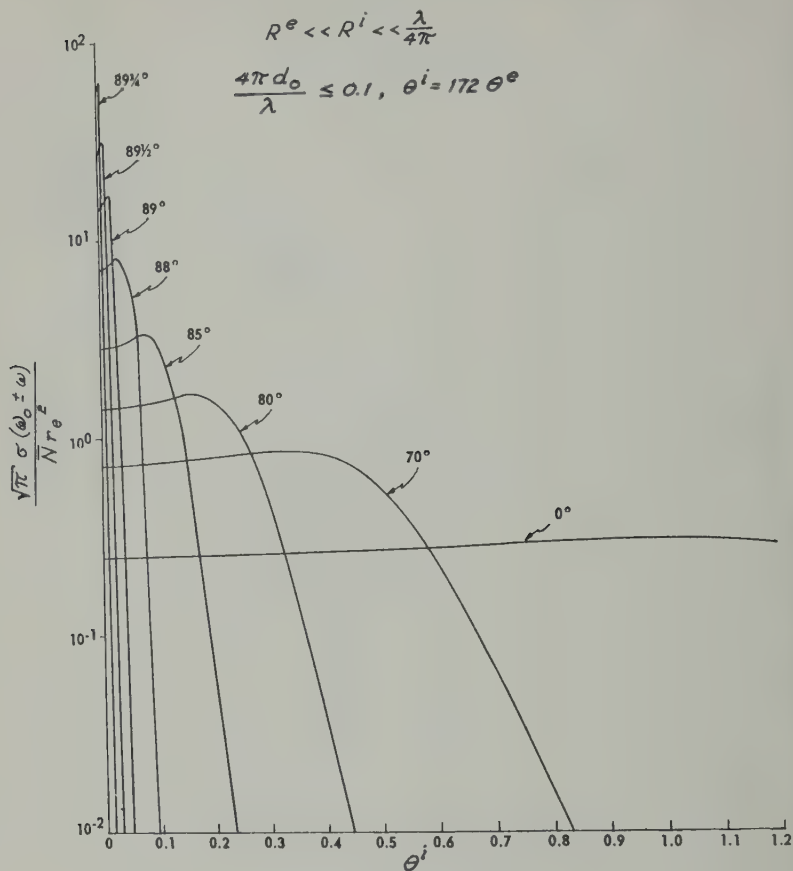


Fig. 15. Spectrum of backscattered echoes.

y_{zz}' yields the same expression for σ_T as equation 56 with $Z = 1$. This check presents further evidence for the validity of the approximations.

5. SUMMARY OF RESULTS

A review of Figures 1 to 16 leads to the following main conclusions for the case of backscattering:

1. For wavelengths much less than the gyroradius of the electron, and consequently much less than the gyroradius of the ions, the spectrum reduces to the spectrum of electromagnetic waves backscattered from an ionized gas in thermal equilibrium in the absence of an applied magnetic field [Dougherty and Farley, 1960; Fejer, 1960; Salpeter, 1960]. Similar results are obtained for any wavelength when the incident plane wave propagation vector is parallel to the magnetic field. On the other hand, as one approaches the

case of propagation perpendicular to the field, all effects described in paragraph 1 below become more pronounced.

2. For wavelengths less than or equal to the electron gyroradius and equal to or smaller than the Debye length, the spectrum envelope, except for normalization, that which is obtained on the basis of electrons as independent scatterers and having a Maxwellian velocity distribution. The spectrum, under the envelope consists of peaks and troughs occurring at harmonics of the gyrofrequency, in agreement with the results of Laaspere [1960].

3. For wavelengths much greater than the electron gyroradius, but smaller than the ion gyroradius, the conclusions for the following cases are: (a) The wavelength is much smaller than the Debye length; for this case the

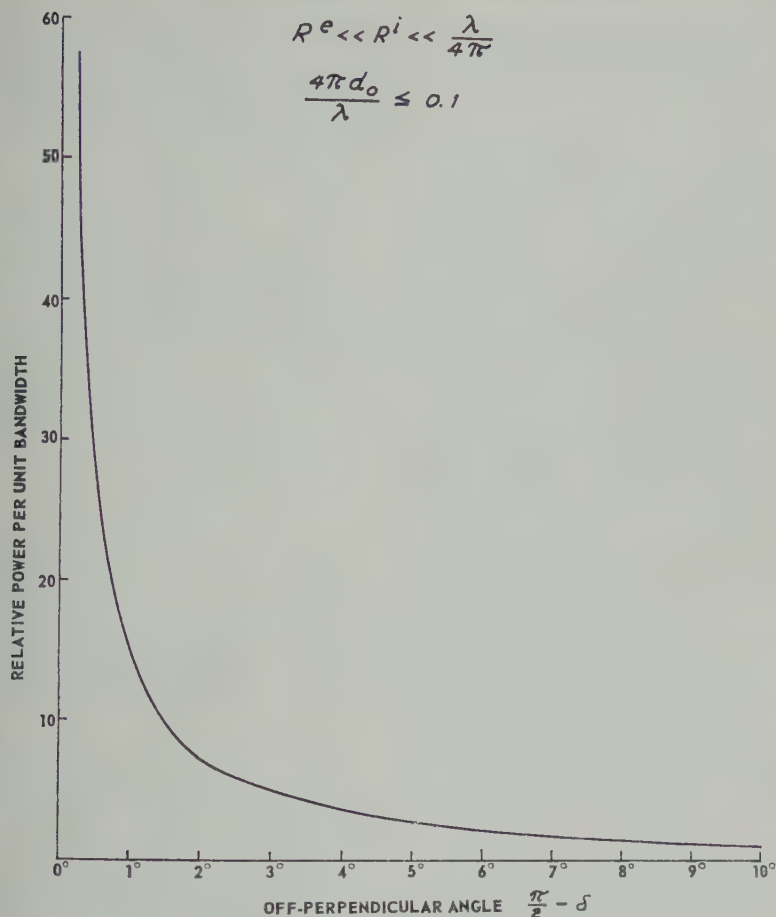


Fig. 16. Relative power as a function of off-perpendicular angle.

which is determined by the ion thermal velocities, is changed but not drastically from that of the isotropic case. (b) The wavelength is less than the Debye length; for this case the spectrum changes rapidly from a low-amplitude, broad spectrum to a large-amplitude, very narrow spectrum as the angle between the magnetic field and the direction of propagation approaches 90° .

For wavelengths of the order of or less than the gyroradius, but much greater than the Debye length, the spectrum function consists of peaks and troughs, with peaks occurring near the harmonics of the gyrofrequency. At low angles, the harmonics of the gyrofrequency are well separated, even though the electrons are the scattering agent. The frequency displacement of the peaks from the harmonics of the ion gyrofrequency decreases as the wavelength decreases

and approaches the Debye length; simultaneously the width of the spectrum increases and approaches the spectrum width determined by the electron thermal velocities.

5. The power per unit bandwidth of the spectra discussed, paragraphs 2 to 4, is large when the direction of propagation is perpendicular to the direction of the magnetic field, and decreases when the propagation direction departs from perpendicularity. These results exhibit some of the characteristics of what ionospheric investigators have called 'field-aligned' or 'aspect-sensitive' echoes. If this theory is used to explain auroral backscatter, the prediction of aspect sensitivity and fading rate are of the right order of magnitude but the calculated intensities are many orders of magnitude less than the measured auroral backscattered intensities. In view of the

discrepancy between the predicted and measured intensities, an 'incoherent scatter' mechanism cannot be the explanation of the auroral backscatter.

6. Plots of the spectrum function for the parameter $0.1 \leq d_0 q \leq 3$ have not been shown; however, calculations for this range indicate that the spectrum changes rapidly and the aforementioned peaks of cases 2 and 4 are not symmetrical about the harmonics of the gyrofrequency and will consist of narrow spikes superimposed upon relatively broader peaks. This asymmetry of the peaks exists for all values of $d_0 q$ less than 3 and is evident in Figure 14.

7. As is shown in section 4, the differential cross section per unit volume is identical with that obtained in the magnetic field-free case.

The peaks and troughs of the spectra seen in the pertinent figures have been emphasized by the use of logarithmic scales for the ordinates. When the ordinates of the spectra are plotted on linear scales the peaks and troughs of the higher harmonics are, in most cases, quite small relative to the amplitude of the spectra about the carrier frequency. Therefore, it may be difficult to observe experimentally the resonances in the spectra due to the applied magnetic field.

APPENDIX A

Derivation of the Relation between the Power Spectrum of the Scattered Field and the Power Spectrum of the Electron Density Fluctuations

When a plane wave with electric intensity $\mathbf{E}_i \exp [-j(\omega_0 t - \mathbf{k}_{\text{inc}} \cdot \mathbf{r})]$ propagates through a dilute plasma, the magnitude of the scattered electric field intensity arising from quasi-static electron density fluctuations $\Delta N(\mathbf{r}, t)$ in a region of volume V is given by

$$E_s = E_i \frac{\exp(jkR_0)}{R_0} r_e \sin \xi \cdot \int_V \Delta N(\mathbf{r}, t) \exp(j\mathbf{q} \cdot \mathbf{r}) dV \exp(-j\omega_0 t) \quad (\text{A1})$$

where R_0 denotes the distance from the coordinate origin within V to the observation point. R_0 is assumed to be large relative to $V^{1/3}$. All other quantities have been defined in the text.

Choose for the region of scatterers a cube of

length L on a side so that $V = L^3$. This condition will cause no loss in generality since that the least dimension of the scattering region is very large compared with the wavelength. Expanding $\Delta N(\mathbf{r}, t)$ in a Fourier series expansion in V , we have

$$\Delta N(\mathbf{r}, t) = \sum_{\mu} \Delta N_{\mathbf{k}_{\mu}}(t) \exp(-j\mathbf{k}_{\mu} \cdot \mathbf{r})$$

The functions $\exp(-j\mathbf{k}_{\mu} \cdot \mathbf{r})$ satisfy periodic boundary conditions on the surface of the cube so that

$$\int_V \exp[j(\mathbf{k}_{\mu} - \mathbf{k}_{\mu'}) \cdot \mathbf{r}] dV = V \delta_{\mu\mu'}$$

From (A3) we conclude

$$V \Delta N_{\mathbf{k}_{\mu}}(t) = \int_V \Delta N(\mathbf{r}, t) \exp[j(\mathbf{k}_{\mu} \cdot \mathbf{r})] dV$$

Since L is very large compared with the wavelength (except for q 's for near forward scattering) choose it so that \mathbf{q} is a member of the set $\{\mathbf{k}_{\mu}\}$. Then

$$V \Delta N_{\mathbf{q}}(t) = \int_V \Delta N(\mathbf{r}, t) \exp(j\mathbf{q} \cdot \mathbf{r}) dV$$

and so (A1) may be written

$$E_s(t) = E_i \frac{\exp(jkR_0)}{R_0} r_e \sin \xi \cdot \int_V \Delta N_{\mathbf{q}}(t) \exp(-j\omega_0 t) dV$$

We will now obtain the relation between the power spectrum of the scattering cross section and the power spectrum of the fluctuations $\Delta N_{\mathbf{q}}(t)$. For this most expeditiously we note: first, the power spectrum of a complex variable is equal to the sum of the power spectra of the real and imaginary parts of the variable; second, the power spectrum of the scattered field intensity may be calculated from either the real or the imaginary part of $E_s(t)$. Thus

$$|\overline{E_s(\omega')}|^2 = 2 |\overline{\text{Re } E_s(\omega')}|^2$$

where the left member denotes the power spectrum of $E_s(t)$ and the right member denotes twice the power spectrum of $\text{Re } E_s(t)$. If quantities such as in (A7) will denote averages over a representative ensemble for the region of interest. Letting σ_T denote the differential

section per unit volume of the scatterers in region of volume V we have

$$V\sigma_T = \frac{R_0^2 \langle [\text{Re } E_s(t)]^2 \rangle}{(|E_i|^2)/2} \quad (\text{A8})$$

where $E_s(t)$ is given by (A6) and $\langle \rangle$ denotes average over all time, i.e.,

$$\langle f(t) \rangle = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T f(t) dt$$

any function $f(t)$. Letting $\sigma(\omega')$ denote the spectrum of σ_T , we have, from (A7) and (A8),

$$\begin{aligned} V\sigma(\omega') &= \frac{R_0^2 |\text{Re } E_s(\omega')|^2}{(|E_i|^2)/2} \\ &= R_0^2 \frac{|E_s(\omega')|^2}{|E_i|^2} \end{aligned} \quad (\text{A9})$$

we can calculate $|E_s(\omega')|^2$ from the autocorrelation of $E_s(t)$ by the Wiener-Khinchine theorem. We have, from (A6),

$$\begin{aligned} \overline{E_s^*(t + \tau) E_s(t)} &= \frac{|E_i|^2}{R_0^2} r_e^2 \\ \sin^2 \xi V^2 \overline{\Delta N_q(t) \Delta N_q^*(t + \tau)} e^{i\omega_0 \tau} & \quad (\text{A10}) \end{aligned}$$

According to the Wiener-Khinchine theorem the Fourier transform of the left member of (A10) is just the spectral density of $E_s(t)$. However, since $\Delta N_q(t)$ is slowly varying relative to $E_s(t)$, the Fourier transform of (A10) will vanish at negative frequencies. Thus the Fourier transform of the left member of (A10) at positive frequencies is just the power spectrum of $E_s(t)$. We have, then,

$$\begin{aligned} \overline{|\tilde{E}_s(\omega)|^2} &= \frac{|E_i|^2}{R_0^2} r_e^2 \sin^2 \xi V^2 \frac{1}{2\pi} \\ &\cdot \int_{-\infty}^{\infty} \overline{\Delta N_q(t) \Delta N_q^*(t + \tau)} e^{i(\omega_0 - \omega')\tau} d\tau \\ &\quad \omega' \geq 0 \end{aligned} \quad (\text{A11})$$

Utilizing (A9),

$$\begin{aligned} \sigma(\omega) &= r_e^2 \sin^2 \xi V \frac{1}{2\pi} \\ &\cdot \int_{-\infty}^{\infty} \overline{\Delta N_q(t) \Delta N_q^*(t + \tau)} e^{-i\omega \tau} d\tau \\ &\quad \omega_0 + \omega \geq 0 \end{aligned} \quad (\text{A12})$$

$$\omega = \omega' - \omega_0.$$

To simplify (A12) we shall show that

$$\text{Im} \overline{\Delta N_q(t) \Delta N_q^*(t + \tau)} = 0 \quad (\text{A13})$$

provided that no external fields are present in the system except a static quasi-uniform magnetic field. Since the representative ensemble of the system is in statistical equilibrium, $\overline{N_q(t) N_q^*(t + \tau)}$ is independent of t . It follows from (A5) that (A13) will be true for arbitrary q and τ if and only if

$$\begin{aligned} \overline{\Delta N(\mathbf{r}, t) \Delta N(\mathbf{r}', t + \tau)} \\ = \overline{\Delta N(\mathbf{r}', t) \Delta N(\mathbf{r}, t + \tau)} \end{aligned} \quad (\text{A14})$$

Since the gas is homogeneous, it is apparent from geometric considerations that (A14) is equivalent to

$$\begin{aligned} \overline{\Delta N_1(\mathbf{r}, t) \Delta N_1(\mathbf{r}', t + \tau)} \\ = \overline{\Delta N_2(\mathbf{r}, t) \Delta N_2(\mathbf{r}', t + \tau)} \end{aligned} \quad (\text{A15})$$

where the left member of (A15) is calculated for one direction of the magnetic field and the right member of (A15) is calculated for the reverse direction of the magnetic field. Now a particular motion of particles which occurs with the magnetic field in one direction cannot occur with the magnetic field in the reverse direction. If we reverse the magnetic field from its original direction, however, the *mirror image* of the *original motion* can occur. In this context, the normal to the mirror is perpendicular to the magnetic field \mathbf{B}_0 .

To prove this, choose rectangular coordinates with the $+z$ axis in the direction of the magnetic field. Then the Hamiltonian equations of motion are invariant to a simultaneous reversal of the magnetic field and reflection of the x axis. The probability of occurrence of a particular microscopic motion with the magnetic field in the $+z$ direction is equal to the probability of occurrence of the motion obtained by reflecting the x axis when the magnetic field is in the $-z$ direction. It follows that an average $\overline{(\Delta N(\mathbf{r}, t) \Delta N(\mathbf{r}', t + \tau))}$ is unchanged by a simultaneous reflection of the x axis and reversal of the \mathbf{B}_0 field. We have then

$$\begin{aligned} \overline{\Delta N_2(x, y, z, t) \Delta N_2(x', y', z', t + \tau)} \\ = \overline{\Delta N_1(-x, y, z, t) \Delta N_1(-x', y', z', t + \tau)} \end{aligned} \quad (\text{A16})$$

Now the transformation $(-x, y, z) \rightarrow (x, y, z)$, $(-x', y', z') \rightarrow (x', y', z')$ is equivalent to a 180° rotation of the points $(-x, y, z)$, $(-x', y', z')$ around the axis of the magnetic field (in the z direction) followed by an equal translation of both points. Since the gas is homogeneous,

$$\overline{\Delta N_1(-x, y, z, t) \Delta N_1(-x', y', z', t + \tau)} \\ = \overline{\Delta N_1(x, y, z, t) \Delta N_1(x', y', z', t + \tau)}$$

Combining this relation with (A16) we obtain (A15), and (A13) follows. The power spectrum of $\Delta N_q(t)$ is therefore

$$|\overline{\Delta N(q, \omega)}|^2 = \frac{2}{2\pi} \cdot \int_{-\infty}^{\infty} \overline{\langle \Delta N_q(t) \Delta N_q^*(t + \tau) \rangle} e^{-i\omega\tau} d\tau \quad (\text{A17})$$

(A12) becomes

$$\sigma(\omega_0 + \omega) = r_e^2 \sin^2 \xi (V/2) \overline{|\Delta N(q, \omega)|^2} \\ \text{for } \omega_0 + \omega \geq 0 \quad (\text{A18})$$

From (A13) and (A17)

$$\overline{|\Delta N(q, \omega)|^2} = \overline{|\Delta N(q, -\omega)|^2} \quad (\text{A19})$$

The normalizations in (A18) are

$$\sigma_T = \int_0^\infty \sigma(\omega') d\omega' \quad (\text{A20})$$

and

$$\overline{\langle |\Delta N_q(t)|^2 \rangle} = \int_0^\infty \overline{|\Delta N(q, \omega)|^2} d\omega \quad (\text{A21})$$

APPENDIX B

Derivations of the Admittance Tensor for an Ionized Gas in Thermal Equilibrium in the Presence of an Externally Applied Uniform Magnetic Field

Two methods of derivation are available, namely, the use of Boltzmann's equation and the particle orbit theory. (A derivation based on particle orbit theory has been made by Dougherty and Farley [draft copy].) In both these methods, the effect of the externally applied fictitious driving force introduces a perturbation in an otherwise Maxwellian distribution. Chandrasekhar [1960] points out that it was Jeans who first demonstrated that, in the absence of col-

lisions, these two methods yield similar results. Although the particle orbit theory is more physical in that it provides better insight into the physical factors involved in a particular problem, Boltzmann's equation provides a more rigorous treatment of complex problems which may not be accessible to the analysis of particle orbit theory.

Bernstein [1958] and others have derived the admittance tensor $Y_{\alpha\beta}$ for a variety of cases. We shall present a derivation (due by Dr. M. Zaidi of this Laboratory) for a single species of gas in the presence of an external magnetic field.

Consider a gas described by a Boltzmann function $f(\mathbf{r}, \mathbf{v}, t)$ whose unperturbed equilibrium is the Maxwellian distribution

$$f_0(\mathbf{v}) = \bar{N} \left(\frac{m}{2\pi KT} \right)^{3/2} \exp \left(-\frac{m\mathbf{v}^2}{2KT} \right)$$

Let a fictitious force be applied, given in the usual notation

$$\mathbf{F} \exp [j(\omega t - \mathbf{q} \cdot \mathbf{r})]$$

The steady-state distribution function is perturbed by the force so that the new Boltzmann distribution function is

$$f_0(\mathbf{v}) + f_1(\mathbf{r}, \mathbf{v}, t)$$

Introducing (B3) in the linearized Boltzmann equation yields

$$\frac{\partial f_1}{\partial t} + \mathbf{v} \cdot \nabla f_1 + \frac{\mathbf{F}}{m} \cdot \frac{\partial f_0}{\partial \mathbf{v}} \\ + \frac{e}{mc} (\mathbf{v} \times \mathbf{B}_0) \cdot \frac{\partial f_1}{\partial \mathbf{v}} = -\nu_c f_1$$

where \mathbf{B}_0 is a constant externally applied magnetic field and

$$-\nu_c f_1 \equiv \left(\frac{\partial f_1}{\partial t} \right)_{\text{collision}}$$

(see Deslodge and Matthysse [1960]); the collision frequency of a charged particle with the neutral particles. (B4) becomes

$$\left(\frac{\partial}{\partial t} + \nu_c + \mathbf{v} \cdot \nabla \right) f_1 - \frac{e}{mc} \mathbf{B}_0 \cdot \mathbf{v} \times \frac{\partial f_0}{\partial \mathbf{v}} \\ = -\frac{\mathbf{F}}{m} \cdot \frac{\partial f_0}{\partial \mathbf{v}}$$

are various methods of solution of equation (1). The simplest is to set

$$\mathbf{v}(\mathbf{r}, t) = f_1(\mathbf{v}) \exp [j(\omega t - \mathbf{q} \cdot \mathbf{r})] \quad (\text{B7})$$

solve for $f_1(\mathbf{v})$.

We introduce a rectangular coordinate system x', y', z' , such that the z' direction is along the z direction. Let the wave vector \mathbf{q} lie in the x', z' plane, making an angle δ with the z' axis, and let the gyrofrequency by $\Omega = (e/mc) |\mathbf{B}_0|$. We use a cylindrical coordinate system in v space such that

$$\mathbf{v} = \hat{e}_1' W \cos \psi + \hat{e}_2' W \sin \psi + \hat{e}_3' U \quad (\text{B8})$$

$$\mathbf{q} = \hat{e}_1' q \sin \delta + \hat{e}_3' q \cos \delta \quad (\text{B9})$$

and \hat{e}_3' are unit vectors along x', y', z' . Then equation B6 may be written

$$-j \left(\frac{\omega'' - \mathbf{q} \cdot \mathbf{v}}{\Omega} \right) f_1(\mathbf{v}) = \frac{\mathbf{F} \cdot \partial f_0}{m \Omega \partial \mathbf{v}} \\ = - \sum_i \frac{F_i v_i f_0}{KT \Omega} \quad (\text{B10})$$

$\omega'' = \omega - j\nu_e$. Equation B10 must be subject to the requirement that $f_1(\mathbf{v})$ be valued, therefore, periodic in ψ with 2π . In order to effect this (see Bernstein) it is convenient to introduce a vector

$$\mathbf{v}' = \hat{e}_1' W \cos \psi' + \hat{e}_2' W \sin \psi' + \hat{e}_3' U$$

function $\Gamma(\psi, \psi')$,

$$\Gamma(\psi, \psi') = \exp \left[j \frac{(\omega'' - q U \cos \delta)}{\Omega} (\psi - \psi') \right. \\ \left. - j q \frac{W \sin \delta}{\Omega} (\sin \psi - \sin \psi') \right]$$

solution of (B10) is

$$f_1(\mathbf{v}) = - \sum_i \frac{F_i f_0(\mathbf{v})}{KT \Omega} \\ \cdot \int_{-\pi}^{\pi} \Gamma(\psi, \psi') v'_i d\psi' \quad (\text{B11})$$

Under a coordinate system in which the components of an arbitrary vector \mathbf{A} transform as spherical harmonics of order

$$A_{\pm 1} = \frac{A_x \pm j A_y}{\sqrt{2}} \quad A_0 = A_z \quad (\text{B12})$$

The scalar product of two vectors \mathbf{A} and \mathbf{B} in this complex coordinate system is defined as

$$\mathbf{A}^* \cdot \mathbf{B} = \sum_{\nu} A_{\nu}^* B_{\nu} = \sum_{\nu} A_{-\nu} B_{\nu} \\ \nu = 1, -1, 0$$

Introducing the new variable

$$\psi'' = \psi - \psi' = -\Omega t$$

equation B11 in the complex coordinate system becomes

$$f_1(\mathbf{v}) = \frac{f_0}{KT} \sum_{\alpha} \int_0^{\infty} F_{\alpha} v_{-\alpha} \exp (-j\alpha \Omega t) \\ \cdot \exp \left[-j(\omega'' - q U \cos \delta) t + j \frac{q W \sin \delta}{\Omega} \right. \\ \left. \cdot \{ \sin \Omega t \cos \psi - \sin \psi (1 - \cos \Omega t) \} \right] dt \quad (\text{B13})$$

$\mathbf{F} \cdot \mathbf{v}'$ has been replaced by

$$\sum_{\alpha} F_{\alpha} v'_{-\alpha} = \sum_{\alpha} F_{\alpha} v_{-\alpha} \exp (-j\alpha \Omega t)$$

in the complex coordinates.

Let

$$S_{x'} = q \sin \delta \sin \Omega t$$

$$S_{y'} = \frac{-q \sin \delta (1 - \cos \Omega t)}{\Omega}$$

$$S_{z'} = q t \cos \delta$$

Therefore

$$f_1(\mathbf{v}) = \frac{f_0}{KT} \sum_{\alpha} \int_0^{\infty} F_{\alpha} v_{-\alpha} \\ \cdot \exp [-j(\alpha \Omega + \omega'') t + j \sum_{\beta} S_{-\beta} v_{\beta}] dt \quad (\text{B14})$$

where the components of \mathbf{S} and \mathbf{q} in the complex coordinate system are

$$S_1 = \frac{S_{x'} + j S_{y'}}{\sqrt{2}} = q_1 \frac{e^{j \Omega t (1)} - 1}{-j \Omega (1)} \\ q_1 = \frac{q \sin \delta}{\sqrt{2}}$$

$$S_{-1} = \frac{S_{x'} - j S_{y'}}{\sqrt{2}} = q_{-1} \frac{e^{-j \Omega t (-1)} - 1}{-j \Omega (-1)}$$

$$q_{-1} = \frac{q \sin \delta}{\sqrt{2}}$$

$$S_0 = S_z = q t \cos \delta = q_0 t$$

$$q_0 = q \cos \delta$$

In general, then,

$$S_\alpha = q_\alpha \left(\frac{e^{-j\Omega t \alpha} - 1}{-j\Omega \alpha} \right) \quad \alpha = 1, -1, 0 \quad (\text{B15})$$

The flux of the particles can now be obtained by integrating the particle velocity over the perturbation, f_1 , in the distribution function. Introducing the single species admittance tensor, Y , defined in connection with (5),

$$S_\alpha = \int_{\mathbf{v} \text{ space}} v_\alpha f_1(\mathbf{v}) d^3 v = \sum_\beta Y_{\alpha\beta} F_\beta \quad (\text{B16})$$

Comparing (B16) and (B14),

$$Y_{\alpha,\beta} = \frac{1}{KT} \int_0^\infty \left\{ \int_{\mathbf{v} \text{ space}} v_\alpha v_{-\beta} f_0(\mathbf{v}) \cdot \exp \left(j \sum_\beta S_{-\beta} v_\beta \right) d^3 \mathbf{v} \right\} \cdot \exp [-j(\beta\Omega + \omega')t] dt \quad (\text{B17})$$

The integral over the velocity space can be evaluated by displacing the origin of the velocity coordinates

$$v_\beta = v_\beta'' + j \frac{KT}{m} S_\beta$$

Then,

$$\begin{aligned} & \int_{\mathbf{v} \text{ space}} v_\alpha v_{-\beta} f_0(\mathbf{v}) \exp \left(j \sum_\beta S_{-\beta} v_\beta \right) d^3 \mathbf{v} \\ &= \int_{\mathbf{v}'', \text{ space}} \left[v_\alpha'' v_{-\beta}'' + 2j \frac{KT}{m} S_\alpha v_{-\beta}'' - \left(\frac{KT}{m} \right)^2 S_\alpha S_{-\beta} \right] \\ & \cdot \exp \left(-\frac{KT}{2m} S^2 \right) f_0(\mathbf{v}'') d^3 \mathbf{v}'' \\ &= \bar{N} \frac{KT}{m} \left(\delta_{\alpha,\beta} - \frac{KT}{m} S_\alpha S_{-\beta} \right) \\ & \cdot \exp \left(-\frac{KT}{2m} S^2 \right) \end{aligned} \quad (\text{B18})$$

where $\delta_{\alpha,\beta}$ is the Kronecker delta, and $S^2 = S^* \cdot S$.

Therefore, the admittance tensor equation B17 reduces to

$$Y_{\alpha\beta} = \frac{\bar{N}}{m} \int_0^\infty \left(\delta_{\alpha,\beta} - \frac{KT}{m} S_\alpha S_{-\beta} \right) \cdot \exp \left[-j(\beta\Omega + \omega')t - \frac{KT}{2m} S^2 \right] dt$$

It will be convenient to introduce the dimensionless variables

$$\theta = (\omega/q)(m/2KT)^{1/2}$$

$$\theta_c = (v_c/q)(m/2KT)^{1/2}$$

and

$$\phi = (\Omega/q)(m/2KT)^{1/2}$$

and the dimensionless tensor y related

$$Y_{\alpha\beta} = \frac{\bar{N}\omega}{KTq^2} y_{\alpha\beta}(\theta, \theta_c, \phi)$$

Thus

$$y_{\alpha\beta} = \frac{1}{2\theta} \int_0^\infty P_{\alpha\beta}(\phi, t) \exp \left[-\theta_c t - \frac{\sin^2 \delta \sin^2 (\phi t/2)}{\phi^2} - \frac{t^2 \cos^2 \delta}{4} \right] dt$$

and $P_{\alpha\beta}$ is defined as

$$P_{\alpha\beta} = \exp [-j(\alpha + \beta)(\phi t/2)] \cdot \left[\delta_{\alpha,\beta} - \frac{2q_\alpha q_{-\beta} \sin \frac{1}{2}\alpha\phi t \sin \frac{1}{2}\beta\phi t}{q^2 \phi^2 \alpha\beta} \right]$$

with α or $\beta = 1, -1, 0$.

The elements of $P_{\alpha\beta}$ are

$$P_{0,0} = 1 - \frac{1}{2} t^2 \cos^2 \delta$$

$$P_{\pm 1, \pm 1} = \exp (\mp j\phi t) \left[1 - \frac{\sin^2 \delta \sin^2 \phi t}{\phi^2} \right]$$

$$P_{\pm 1, \mp 1} = -\frac{\sin^2 \delta \sin^2 (\phi t/2)}{\phi^2}$$

$$P_{\pm 1, 0} = P_{0, \pm 1} = \exp \left(\mp j \frac{\phi t}{2} \right)$$

$$\cdot \left[\frac{\cos \delta \sin \delta}{\phi \sqrt{2}} t \sin \frac{\phi t}{2} \right]$$

For particles with negative charge replace $-\phi$ in equations B22 and B23. Note that in equation B21 the collisional frequency assures the convergence of the integrals for values of the other parameters. If we set

we have done in the text, the integral converges as long as $\delta < \pi/2$.

etermination of the Admittance Tensor Elements in a Coordinate System with \mathbf{q} in the z Axis Direction

The next task is the calculation of the quantity $\mathbf{Y} \cdot \mathbf{q}$ of equation 13. This calculation will be facilitated by the choice of a coordinate system with the z axis in the direction of \mathbf{q} so that

$$\mathbf{q} \cdot \mathbf{Y} \cdot \mathbf{q} = q^2 Y'_{zz} \quad (\text{B24})$$

such a system is readily obtained by a rotation of the axes of the x', y', z' system through an angle δ around the y' axis. This leads to a system, x, y, z , defined by

$$\begin{aligned} x &= x' \cos \delta - z' \sin \delta \\ y &= y' \\ z &= x' \sin \delta + z' \cos \delta \end{aligned} \quad (\text{B25})$$

By combining the transformation in (B25) with that in (B12), the over-all transformation from the complex system to the x, y, z system can be obtained. Letting $\hat{e}_1, \hat{e}_2, \hat{e}_3$ denote the unit vectors in the x, y, z system, their components in the complex system are

$$\begin{aligned} \hat{e}_1 &= \left(\frac{\cos \delta}{\sqrt{2}}, \frac{\cos \delta}{\sqrt{2}}, -\sin \delta \right) \\ \hat{e}_2 &= \left(\frac{j}{\sqrt{2}}, \frac{-j}{\sqrt{2}}, 0 \right) \\ \hat{e}_3 &= \left(\frac{\sin \delta}{\sqrt{2}}, \frac{\sin \delta}{\sqrt{2}}, \cos \delta \right) \end{aligned} \quad (\text{B26})$$

The order of components in the complex system is 1, -1, 0.

The elements y_{ij} of \mathbf{y} in the x, y, z system are conveniently obtained by transforming the elements $P_{\alpha\beta}$ to obtain the elements P_{ij} in the x, y, z system and then employing the relation

$$y_{\alpha\beta} = \frac{1}{2\theta} \int_0^\infty P_{\alpha\beta} e^{-j\theta t + \gamma} dt \quad (\text{B27})$$

$$\frac{\sin^2 \delta}{\phi^2} \sin^2 \frac{1}{2} \phi t - \frac{1}{4} t^2 \cos^2 \delta \quad (\text{B28})$$

It follows from (B21) when we place $\theta_0 = 0$.

The elements P_{ij} can be obtained from the $P_{\alpha\beta}$ by the usual tensor transformation. This may be written

$$P_{ij} = \sum_{\alpha, \beta} \hat{e}_{i\alpha}^* P_{\alpha\beta} \hat{e}_{j\beta} \quad (\text{B29})$$

where $\hat{e}_{i\alpha}$ denotes the α component in the complex system of the unit vector \hat{e}_i in the x, y, z system. The particular form of the transformation from the $P_{\alpha\beta}$ to the P_{ij} in (B29) is obtained by first noting that, since the transformation from the complex system to the x, y, z system is unitary, $\hat{e}_i^* \cdot \mathbf{P} \cdot \hat{e}_i$ is an invariant scalar. Then, evaluating $\hat{e}_i^* \cdot \mathbf{P} \cdot \hat{e}_i$ first in the x, y, z system and then in the complex system, we obtain (B29).

The transformation in (B29) is most conveniently performed in two steps by putting

$$P_{ij} = \sum_{\alpha} \hat{e}_{i\alpha}^* P_{\alpha j} \quad (\text{B30})$$

where

$$P_{\alpha j} = \sum_{\beta} P_{\alpha\beta} \hat{e}_{j\beta} \quad (\text{B31})$$

Utilizing the expressions for the elements $P_{\alpha\beta}$ given by (B23) and noting that the $\hat{e}_{j\beta}$ are given by (B26), we obtain for $P_{\alpha\beta}$, after considerable manipulation,

$$\left. \begin{aligned} P_{0x} &= \frac{t \cos^2 \delta}{\sin \delta} \left\{ \frac{d\gamma}{dt} + \frac{1}{2} t \right\} - \sin \delta \\ P_{*1x} &= \frac{\cos \delta}{\sqrt{2}} \left\{ -h^2 + e^{\mp i\phi t} (1 - h^2) \right. \\ &\quad \left. + \frac{t \cot \delta}{\sqrt{2}} \left[\frac{dh^2}{dt} \mp \phi h^2 \right] \right\} \\ P_{0y} &= -th^2 \phi \cot \delta \\ P_{*1y} &= \pm \frac{j}{\sqrt{2}} \{ h^2 + e^{\mp i\phi t} (1 - h^2) \} \\ P_{0z} &= \cos \delta e^{-\gamma} \frac{d}{dt} [te^{\gamma}] \\ P_{*1z} &= \pm \frac{\sin \delta}{\sqrt{2}} e^{-\gamma} \frac{d}{dt} \left[\frac{1 - e^{\mp i\phi t}}{j\phi} e^{\gamma} \right] \end{aligned} \right\} \quad (\text{B32})$$

and

$$h = \frac{\sin \delta}{\phi} \sin \frac{1}{2} \phi t \quad (\text{B33})$$

Now the P_{ij} can be evaluated by (B30), and then the y_{ij} can be obtained from (B27).

As will be shown in Appendix C, only y_{zz} will be needed to obtain a good approximation to y_{zz}' (equation 17) in the cases of current interest. The criterion which will be obtained for the validity of this approximation does not depend on the form of the elements of \mathbf{y} but only on their upper bound. Thus, owing to the cumbersome character of the expressions for y_{ij} we shall only list y_{zz} and y_{zz}' , the latter being typical of the forms of the remaining elements.

$$y_{zz} = j + \theta \int_0^\infty e^{-i\theta t + \gamma} dt \quad (\text{B34})$$

$$\begin{aligned} y_{zz}' &= j \tan^2 \delta + \theta \tan^2 \delta \int_0^\infty e^{-i\theta t + \gamma} dt \\ &\quad - \frac{j \tan^2 \delta}{\phi} \int_0^\infty e^{-i\theta t + \gamma} \sin \phi t dt \\ &\quad + \frac{\sec^2 \delta}{\theta} \int_0^\infty e^{-i\theta t + \gamma} \cos \phi t dt \\ &\quad - \frac{\tan^2 \delta}{4\theta\phi^2} \int_0^\infty e^{-i\theta t + \gamma} \sin^2 \phi t dt \quad (\text{B35}) \end{aligned}$$

A careful study of all elements y_{ij} reveals that

$$|y_{ij}| < \frac{\sec \delta}{\theta} \quad \text{all } i, j \quad (\text{B36})$$

for the cases of large ϕ (strong magnetic field) and small ϕ (weak magnetic field). It is taken for granted in this paper that (1) will also be satisfied for intermediate values of ϕ . Equation (B36) will prove useful in the discussion of Appendix C.

APPENDIX C

Simplification of the Effective Admittance Tensor \mathbf{Y}'

As shown in the text, equation 17:

$$\mathbf{y}' = (\mathbf{y}^i - \mathbf{g}) \cdot (\mathbf{y}^i + \mathbf{y}^e - \mathbf{g})^{-1} \cdot \mathbf{y}^e = \mathbf{A} \cdot \mathbf{B} \cdot \mathbf{C}$$

(note that \mathbf{A} , \mathbf{B} , and \mathbf{C} are algebraic quantities and should not be confused with quantities like B , the fluctuating magnetic field, used in text), where

$$\mathbf{A} = (\mathbf{y}^i - \mathbf{g})$$

$$\mathbf{B} = (\mathbf{y}^i + \mathbf{y}^e - \mathbf{g})^{-1} = \mathbf{D}$$

$$\mathbf{C} = \mathbf{y}^e$$

It will be convenient to write $\mathbf{y}^e = \mathbf{y}^i + \mathbf{y}^e - \mathbf{g}$ and the elements of \mathbf{y}^i and \mathbf{y}^e are listed in Appendix B, with the appropriate subscripts. Moreover,

$$\mathbf{B} = \begin{bmatrix} + \begin{vmatrix} D_{22} - D_{23} \\ D_{32} - D_{33} \end{vmatrix} & - \begin{vmatrix} D_{12} D_{13} \\ D_{32} D_{33} \end{vmatrix} & + \begin{vmatrix} D_{12} D_{13} \\ D_{22} D_{23} \end{vmatrix} \\ - \begin{vmatrix} D_{21} D_{23} \\ D_{31} D_{33} \end{vmatrix} & + \begin{vmatrix} D_{11} D_{13} \\ D_{31} D_{33} \end{vmatrix} & - \begin{vmatrix} D_{11} D_{13} \\ D_{21} D_{23} \end{vmatrix} \\ + \begin{vmatrix} D_{21} D_{22} \\ D_{31} D_{32} \end{vmatrix} & - \begin{vmatrix} D_{11} D_{12} \\ D_{31} D_{32} \end{vmatrix} & + \begin{vmatrix} D_{11} D_{12} \\ D_{21} D_{22} \end{vmatrix} \end{bmatrix}$$

$$\begin{vmatrix} D_{11} D_{12} D_{13} \\ D_{21} D_{22} D_{23} \\ D_{31} D_{32} D_{33} \end{vmatrix}$$

where $|D|$ in the denominator is the determinant of the \mathbf{D} matrix. The components of \mathbf{D} are

$$\begin{aligned} D_{11} &= y_{zz}^e - jQ & D_{12} &= y_{xy}^e & D_{13} &= y_{xz}^e \\ D_{21} &= y_{yz}^e & D_{22} &= y_{yy}^e - jQ & D_{23} &= y_{yz}^e \\ D_{31} &= y_{zx}^e & D_{32} &= y_{xy}^e & D_{33} &= y_{zz}^e + j d_0^2 q^2 \end{aligned}$$

$$Q = d_0^2 q^2 c^2 q^2 / \omega^2$$

Now the determinant common to all the elements of B is

$$D = D_{11} D_{22} D_{33} - D_{11} D_{23}^2 - D_{12}^2 D_{33} + 2 D_{12} D_{23} D_{13} - D_{13}^2 D_{22} \quad (C3)$$

$$D = D_{11} D_{22} \left[D_{33} - \frac{D_{23}^2}{D_{22}} - \frac{D_{12}^2 D_{33}}{D_{11} D_{22}} + \frac{2 D_{12} D_{23} D_{13}}{D_{11} D_{22}} - \frac{D_{13}^2}{D_{11}} \right]$$

$Q \gg |y_{ii}^*|$. Then $D_{11} \approx D_{22} \approx -jQ$.

define

$$\beta = j \frac{D_{ij}}{Q} \quad i \neq j$$

Real $\beta \ll 1$ and $\text{Im } \beta \ll 1$.

$$D = D_{11} D_{22} [D_{33} - \beta D_{23} - \beta^2 D_{33} + 2\beta^2 D_{12} D_{23} D_{13} - \beta D_{13}^2] \approx D_{11} D_{22} D_{33} \quad (C4)$$

where

$$\frac{D_{11} D_{22} - D_{21}^2}{\det D} \approx \frac{D_{11} D_{22}}{D_{11} D_{22} D_{33}} = \frac{1}{D_{33}} \quad (C5)$$

D_{11}^2 has been neglected since it is of the order of $D_{11} D_{22} / \beta^2$. Similarly,

$$\frac{1}{D_{22}} = \frac{-1}{jQ} \text{ and } B_{zz} = \frac{1}{D_{11}} = \frac{-1}{jQ} \quad (C6)$$

Off diagonal components consist of

$$B_{xi} = \frac{D_{21} D_{32} - D_{31} D_{22}}{D_{11} D_{22} D_{33}}$$

$$\approx -B_{zz}(-\beta^2 + j\beta)$$

$$B_{yz} = \frac{D_{31} D_{23} - D_{21} D_{33}}{D_{11} D_{22} D_{33}}$$

$$\approx -\beta^2 B_{zz} + j \frac{\beta}{Q} = -\beta^2 B_{zz} + \beta B_{yz}$$

$$B_{xy} = \frac{D_{31} D_{12} - D_{11} D_{32}}{D_{11} D_{22} D_{33}}$$

$$\approx B_{zz}(-\beta^2 + j\beta)$$

$$B \approx \begin{bmatrix} j/Q & \beta B_{yz} & jB_{zz}\beta \\ \beta B_{yz} & j/Q & jB_{zz}\beta \\ jB_{zz}\beta & jB_{zz}\beta & B_{zz} \end{bmatrix}$$

$$B \approx \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & B_{zz} \end{bmatrix}$$

+ (terms of order $|B_{ij}|/Q$ or smaller) (C7)

Therefore, provided that

$$Q = \frac{d_0^2 q^2 c^2 q^2}{\omega^2} \gg |y_{ii}^* + y_{ii}^i|$$

where the subscripts i, j refer to components of the admittance tensors, the second matrix of equation C7 may be neglected compared to the first one, and using equation C5 for B_{zz} we obtain

$$B \approx \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \frac{1}{y_{zz}^* + y_{zz}^i + j d_0^2 q^2} \end{bmatrix} \quad (C8)$$

According to equation C1, and since B has only a zz component,

$$y'_{zz} = A_{zz} B_{zz} C_{zz} = \frac{(y_{zz}^i + j d_0^2 q^2) y_{zz}^*}{y_{zz}^i + y_{zz}^* + j d_0^2 q^2} \quad (C9)$$

The upper bounds of the admittance tensors are, from (B36),

$$|y_{ii}^* + y_{ii}^i|_{\max} \leq \frac{\sec \delta}{\theta^*} + \frac{\sec \delta}{\theta^i} \approx \frac{\sec \delta}{\theta^*}$$

since $\theta^i \gg \theta^*$. The condition for using only the zz terms can then be expressed as

$$\frac{d_0^2 q^2 c^2 q^2}{\omega^2} \gg \frac{\sec \delta}{\theta^*}$$

Using the definition of θ^* , the criterion reduces to

$$d_0^2 q^2 \frac{c q}{\omega} \left(\frac{m^* c^2}{2KT} \right)^{1/2} \gg \sec \delta \quad (C10)$$

or using the definitions for d_0 , θ^* , and q

$$2 \sin^2 \left(\frac{\epsilon}{2} \right) \left(\frac{f_0}{f_{pe}} \right)^2 \gg \sec \delta \left[\frac{\omega}{2\omega_0 \sin(\epsilon/2)} \left(\frac{m^* c^2}{2KT} \right)^{1/2} \right] \quad (C11)$$

where f_{pe} is the electron plasma frequency and ϵ

is the angle between \mathbf{k}_{inc} and \mathbf{k} . The bracketed quantity is, by definition, θ^* . As ϵ approaches zero, the condition for forward scattering, it becomes more and more difficult to satisfy the criterion C11. However, for $\pi/3 < \epsilon \leq \pi$, $\sin(\epsilon/2) \geq \frac{1}{2}$ so that (C11) will be satisfied when

$$\left(\frac{f_0}{f_{pe}}\right)^2 \gg (\sec \delta) \theta^* \left[\frac{m^*}{m^i}\right]^{1/2} \quad (\text{C12})$$

The spectrum calculations in the main body of the report show that, for $d_0 q < 0.1$ ($\lambda > d_0$), the spectrum is reduced to 0.01 of its maximum value for $\theta^* \approx 3$, so that we may replace θ^* by 3. The heavier the ions, the easier it is to satisfy criterion C12. Since criterion C10 is satisfied for $d_0 q$ much less than 1 ($\lambda > d_0$), it is obviously satisfied when $d_0 q$ is much greater than 1 ($\lambda < d_0$).

For ionospheric backscattering [$\sin(\epsilon/2) = 1$] applications, when the plasma frequency is taken as high as 10 Mc/s and the operating frequency as low as 50 Mc/s, our approximations are still valid over the range

$0 \leq \delta \leq 89.4^\circ$ when protons are the dominant ion species

$0 \leq \delta \leq 89.8^\circ$ when the mass of the ions is that of oxygen atom

Very High Frequency Approximation

When

$$d_0^2 q^2 \gg |y_{zz}^i + y_{zz}^e|$$

we note that equation C9 reduces to

$$y'_{zz} = y_{zz}^e \quad (\text{C13})$$

From our study on the upper bounds of y_{zz}^e and y_{zz}^i , we find that (C13) is a valid expression when $d_0 q \geq 10$.

Equation C13 demonstrates that, for wavelength of the order of or less than the Debye length, the backscattered power spectrum is as though each electron scattered 'incoherently' without any Coulomb interaction with the positive ions.

Acknowledgments. We are indebted to Dr. M. Zaidi for helpful discussions and suggestions, and we wish to thank Professor Salpeter and Drs. Dougherty and Farley for sending us a draft copy of their treatment of this problem while this paper was in preparation. Their approximation to the effective admittance tensor advanced the final solution of our work, and we are greatly indebted to them.

This work was supported by the intersearch program of Cornell Aeronautical Laboratory.

Note added in revised paper. We wish to thank Drs. J. Fejer's, T. Hagfors', and E. Salpeter's papers have now been published.

REFERENCES

- Bernstein, I. B., *Phys. Rev.*, **109**, 10, 1958.
 Bowles, K. L., *Natl. Bur. Standards Rep.* 607.
 Callen, H. B., and T. A. Welton, *Phys. Rev.*, **101**, 34-40, 1951.
 Callen, H. B., and R. F. Greeno, *Phys. Rev.*, **101**, 702-711, 1952.
 Callen, Barasch, and Jackson, *Phys. Rev.*, **101**, 1382-1386, 1952.
 Chandrasekhar, S., *Plasma Physics*, p. 13, University of Chicago Press, Chicago, 1960.
 deGroot, S. R., *Thermodynamics of Irreversible Processes*, Interscience Publishers, New York, 1951.
 Desloge, E. A., and S. W. Matthyssse, *Phys.*, **28**, 1-11, 1960.
 Dougherty, J. P., and D. T. Farley, *Proc. Roy. Soc. London, A*, **259**, 77-99, 1960.
 Dwight, H. B., *Tables of Integrals and Other Mathematical Data*, The Macmillan Co., New York, 1955.
 Ekstein, H., and N. Rostoker, *Phys. Rev.*, **101**, 1023-1028, 1955.
 Fejer, J. A., *Can. J. Phys.*, **38**, 1114-1123, 1960.
 Fejer, J. A., *Can. J. Phys.*, **39**, 716-740, 1961.
 Greene, R. F., and H. B. Callen, *Phys. Rev.*, **101**, 1387-1391, 1952.
 Hagfors, T., *J. Geophys. Research*, **66**, 1661-1664, 1961.
 Jackson, J., *Phys. Rev.*, **87**, 471-472, 1952.
 Kaplan, T. A., *Phys. Rev.*, **102**, 1447-1451, 1956.
 Kubo, R., and K. Tomita, *J. Phys. Soc. Japan*, **10**, 888, 1954.
 Kubo, R., *Can. J. Phys.*, **34**, 1274, 1956.
 Kubo, R., *J. Soc. Phys. Japan*, **12**, 570, 1955.
 Laaspere, T., Ph.D. thesis, Cornell University, September 1960.
 Lax, M., *Phys. Rev.*, **109**, 1921-1926, 1958.
 Lax, M., *Revs. Modern Phys.*, **32**, 25-64, 1960.
 Onsager, L., *Phys. Rev.*, **37**, 405, 38, 2265, 1960.
 Pineo, V. C., L. G. Kraft, and H. W. Briscoe, characteristics of ionospheric backscattered at 440 Mc/s, *J. Geophys. Research*, **66**, 2629-2633, 1960.
 Renau, J., Scattering of electromagnetic waves from a nondegenerate ionized gas, *J. Geophys. Research*, **66**(11), 3631-3640, 1960.
 Richardson, J., *Inst. Radio Engrs. Trans. on Theory IT-1*, No. 1 pp. 62-65, 1955.
 Salpeter, E. E., *Phys. Rev.*, **120**, 1528-1532, 1960.
 Salpeter, E. E., Scattering of radio waves from ions above the ionosphere, *J. Geophys. Research*, **65**(6), 1851-1852, 1960b.
 Salpeter, E. E., *Phys. Rev.*, **122**, 1663, 1960c.

(Manuscript received April 7, 1961; revised May 29, 1961.)

Prediction of VLF Diurnal Phase Changes and Solar Flare Effects

W. D. WESTFALL

U. S. Navy Electronics Laboratory, San Diego 52, California

Abstract. VLF diurnal phase shift data are compared with predictions resulting from the waveguide equations for the first order mode. The effects of higher order mode energies appear to exist out to 2700 km, and may at times affect measurements beyond this range. A relation is given for the prediction of diurnal phase shift for 10–20 kc radio waves. Expressions are given for relating observed VLF phase perturbations caused by both small and large solar flares to the resulting apparent decreases in ionospheric reflecting height.

During 1960 extensive phase comparison measurements have been made at U. S. Navy Electronics Laboratory facilities at very low frequencies. In addition to recording the phase of the transmissions from the U. S. Navy stations at Panama and Hawaii (NBA on 18 kc and WWVL on 19.8 kc) and WWVL on 20 kc (National Bureau of Standards), the U. S. Navy Electronics Laboratory Radux-Omega navigation transmitters operating on frequencies between 9.2 and 15.2 kc from their facilities at Honolulu, Hawaii, and Forestport, New York, have been monitored.

Table 1 shows (a) diurnal phase change data which have previously been published elsewhere (Crombie, Bracewell, Straker, and Westcott, 1952; Crombie, Allan, and Newman, 1958; Pierce, 1957; Silkwood, 1959) and (b) phase measurements of NBA, NPM, and WWVL taken specially for this study of VLF phase fluctuations. (c) data taken on several frequencies and extracted from the Radux-Omega measurement program.

The data were taken on frequencies between 10 and 20 kc with path distances between 535 and 1000 km. Paths completely over land and completely over sea have been used, as well as paths consisting of both land and sea. The season of the year during which the measurements were made has not been considered.

For the first two paths, Crombie, Allan, and Newman [1958] and Pierce [1957] have fixed values of 70 ± 2 and 34 ± 1 microseconds respectively. The new data represent the average of at least several magnetically quiet days, and the repeatability of any measurement for any one path is a function of the variability of the iono-

sphere. Where not previously published, standard deviations for the data are shown in the last column in Table 1.

Wait [1959] has pointed out that the phase velocity of a mode of order n in a waveguide consisting of the earth and a concentric sharply bounded and isotropic ionosphere is approximated by

$$V = c(1 - C_n^2)^{-1/2} \left(1 - \frac{h}{2a}\right) \quad (1)$$

where

c = velocity of light.

$C_n = (n - \frac{1}{2}) \lambda / 2h$.

h = height of the ionospheric reflecting layer.

a = radius of the earth.

λ = wavelength.

Equation 1, written for the principal mode, takes the form

$$V = c \left(1 - \frac{\lambda^2}{16h^2}\right)^{-1/2} \left(1 - \frac{h}{2a}\right) \quad (2)$$

The diurnal phase shift measured at a particular distance from the transmitting station can be expressed by

$$\Delta t = - \left(\frac{d}{V_1} - \frac{d}{V_2} \right) = - \frac{d}{c} \left(\frac{c}{V_1} - \frac{c}{V_2} \right) \quad (3)$$

where d is the distance from the transmitter and Δt is the phase change in seconds resulting from the difference in day and night phase velocities V_1 and V_2 of the principal mode.

Designating the daytime and nighttime heights as h and $(h + \Delta h)$ respectively, the expression

TABLE 1. Diurnal Phase Shifts for 10 to 20 Kc Radio Waves Over Path Distances of 535 to 11000 km

Frequency, kc/s	Path and Distance, km	Path Composi- tion, % land	Diurnal Phase Shift, micro-sec.	Standard Deviation, micro-sec.
16	Rugby-New Zealand ¹ 11000	...	70	...
16	Rugby-Cambridge ² 5200	0	34	...
16	Rugby ³ 535	100	35	...
18.6	Seattle-San Diego ⁴ 1780	100	18	...
10.2*	Haiku-San Diego 4180	0	52	2.6
19.8†	Haiku-San Diego 4180	0	27	3.0
18†	Panama-San Diego 4670	90	27	4.7
10.2*	Haiku-Boulder 5360	25	61	1.7
10.2*	Haiku-Forestport, N. Y. 7833	50	87	5.6
18†	Panama-Haiku 8417	0	50	6.4
18†	Panama-Forestport, N. Y. 3851	0	29	1.2
20†	Boulder, Col.-San Diego 1400	100	17	3.1

* Extracted from USNEL Radux-Omega data.
† USNEL phase shift measurement.
¹ Crombie, Allen, and Newman.
² Pierce.
³ Bain and co-workers.
⁴ Silkwood.

for Δt is

$$\Delta t = -\frac{d}{c} \left\{ \frac{\left[1 - \frac{\lambda^2}{16h^2} \right]^{1/2}}{\left[1 - \frac{h}{2a} \right]} - \frac{\left[1 - \frac{\lambda^2}{16(h + \Delta h)^2} \right]^{1/2}}{\left[1 - \frac{h + \Delta h}{2a} \right]} \right\} \quad (4)$$

for paths long enough that only one mode need be considered.

When both $\lambda^2/16h^2$ and $h/2a$ are small compared to unity, expression (4) can be approximated by

$$\frac{c \Delta t}{d} = \frac{\Delta h}{2a} + \frac{\lambda^2 A}{32} \quad (5)$$

where

$$A = \frac{1}{h^2} - \frac{1}{(h + \Delta h)^2}$$

The data from the five longest paths in Table 1 were used in (5) to get a least squares estimate of A and Δh . The corresponding value of h was obtained by solving (6) iteratively. From the data the daytime

TABLE 2. Comparison of Measured and Calculated Diurnal Phase Shifts Over Long Paths

Distance, km	f , kc	μsec , Measured	Calculated
11000	16	70	70
8417	18	50	50
7833	10.2	87	87
5360	10.2	61	61
5200	16	34	34

is 59.8 km and the diurnal height change is 5 km. These values do not represent the actual heights of the ionosphere but are effective values which are to be used in predicting VLF diurnal phase changes.

The diurnal phase change Δt can be predicted along paths as a function of distance and wavelength by rewriting (4) as

$$-\frac{\Delta h}{c} d \left\{ \left(\frac{\lambda^2}{16h^3} + \frac{1}{2a} \right) - \frac{\lambda^2}{32h^2} \left[\frac{3}{h^2} \frac{\Delta h}{h^2} - \frac{4(\Delta h)^2}{h^3} \right] \right\} \quad (7)$$

Substituting the estimated values of 59.8 and 5 km in (7) gives

$$= (9.4 \times 10^{-6} \lambda^2 + 3.2 \times 10^{-3}) d = F(d, \lambda) \quad (7a)$$

Δt is in microseconds and d and λ are in meters. Table (2) shows the comparison of diurnal changes predicted by (7a) with the observed changes for the five long paths.

For shorter paths, the diurnal phase change is given by

$$|\Delta t| = F(d, \lambda) m(d) \quad (8)$$

$m(d)$ is an empirical multiplying factor accounting for short paths the effects of higher order modes. Figure 1 shows $m(d)$ calculated

from the possible multimode paths of Table 1.

For distances between 5000 and 2000 km, the curve in Figure 1 is drawn as a dual and dashed line because until further experimental data become available the effects of higher order modes are uncertain. At short distances the diurnal phase shift may also exhibit seasonal changes as well as a greater day-to-day variability from that observed at the longer distances.

By using (8) and the value of $m(d)$ from Figure 1, predictions of diurnal phase change can be made for VLF propagation as a function of distance and wavelength.

Expression (7) relates any observed phase change Δt to a corresponding height change Δh for a path where only one mode need be considered. For height changes of the magnitude of the diurnal shift the ratio $\Delta h/h$ is not small compared to unity and the last two terms on the right-hand side of (7) are required for accuracy. However when (7) and (8) are used to relate the observed phase variations resulting from small solar flares to corresponding perturbations in ionospheric height, the ratio $\Delta h/h$ is small compared to unity.

For the case of an ionospheric perturbation caused by a solar flare the height fluctuation is superimposed on the daytime height. The factor $m(d)$ which contains the effects of higher order modes on diurnal shift is determined more by the nighttime than daytime conditions. The

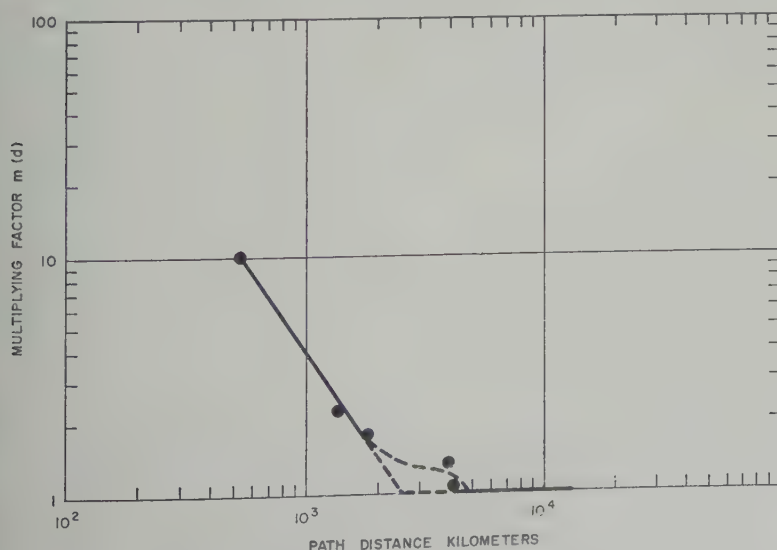


Fig. 1. Multiplying factor $m(d)$ as a function of path distance.

corresponding factor for flares is equal to unity for distances beyond 1000 km and appears to be smaller than $m(d)$ for shorter distances.

For paths longer than 1000 km the relation between the observed phase change due to a solar flare and the resulting height change is given by

$$\Delta h = - \frac{c \Delta t}{d \left[\frac{\lambda^2}{16h^3} + \frac{1}{2a} \right]} \quad \left(\begin{array}{l} \text{small} \\ \text{flares} \end{array} \right) \quad (9)$$

where h is now the height just before the flare perturbation occurs. For flares indicated by phase changes of the order of 10 per cent of the diurnal change or less, equation (9) can be used to estimate the perturbation in ionospheric reflection height with a precision that is consistent with that of the measurements.

The height change resulting from a large flare can be obtained by modifying (9) to include the third term of (7) so that

$$\Delta h = - \frac{c \Delta t}{d \left(\frac{\lambda^2}{16h^3} + \frac{1}{2a} \right)} \cdot \left[1 + \frac{3\lambda^2 c \Delta t}{32h^4 d \left(\frac{\lambda^2}{16h^3} + \frac{1}{2a} \right)^2} \right] \quad \left(\begin{array}{l} \text{large} \\ \text{flares} \end{array} \right) \quad (10)$$

Expression 10 provides an estimate of Δh with an approximation error of less than 2 per cent

even for flares as large as half the diurnal

Acknowledgments. This work was supported by the Advanced Research Project Agency, ARPA order No. 164-61.

¹ Height changes between 60 and 100 MHz induced by solar flares can also be determined from long paths from a set of curves by Wait and [1961]. These curves show the absolute phase velocity, based on improved wave theory, for 8- to 30-kc radio waves. Equal and an ambient height of 70 km gives changes which are in quite good agreement with these curves over the 10- to 20-kc range.

REFERENCES

- Bain, W. C., R. N. Bracewell, T. W. Straker, C. H. Westcott, The ionospheric properties of radio waves of frequency 16 kc/s at distances of about 540 km, *Proc. IEE*, 99, 1952.
- Crombie, D. D., A. H. Allan, and M. M. Phase variations of 16 kc/s transmissions from Rugby as received in New Zealand, *Proc. IRE*, B, 301-304, May 1958.
- Pierce, J. A., Intercontinental frequency comparison by very low-frequency radio transmissions, *Proc. IRE*, 45, 704-803, June 1957.
- Silkwood, J. A., Phase stability measurements of 18.6 kc radio transmissions, *U. S. Naval Electronics Lab. Rept. 899*, June 1959.
- Wait, J. R., Diurnal change of ionospheric reflection height deduced from phase velocity measurements at VLF, *Proc. IRE*, 47, 589, 1959.
- Wait, J. R., and Kenneth P. Spies, A study of the phase velocity of VLF radio waves, *J. Geophysical Research*, 66, 992-993, 1961.

(Manuscript received May 15, 1961; revised June 22, 1961.)

VLF Ionospheric Reflection Coefficients—Derivation from Impedance Concepts and Values for Some Model Ionospheres

EDWARD C. FIELD¹ AND PAUL TAMARKIN

*Electronics Department, The Rand Corporation
Santa Monica, California*

Abstract. A convenient general method is presented for computing the reflected wave excited by an arbitrarily polarized, plane, VLF electromagnetic wave incident at an arbitrary direction upon a sharply bounded, absorptive ionosphere in the geomagnetic field. The relative simplicity of the method results from the use of impedance concepts coupled with Booker's treatment of the propagation vector. Expressions for reflection coefficients are given that are exact (for the model assumed) in the case of propagation perpendicular to magnetic meridians and indicate clearly lack of east-west and west-east reciprocity. The results of calculations for eight model ionospheres are given in the form of graphs of reflection coefficient amplitude and phase versus angle of incidence for both east-west and west-east propagation at 16 kc/sec. The model ionospheres are characterized by values of electron density and collision frequency. A brief relation is made between the computations and some experimental results.

INTRODUCTION

The treatment of the oblique reflection of radio waves from a sharply bounded ionosphere in the geomagnetic field has been a problem of interest for many years. We will be concerned here with the reflection of very low frequency waves over the frequency range for which the assumption that the ionosphere is sharply bounded provides a reasonable working model. (We do not imply that the ionospheric wave is actually returned to earth by reflection from a sharp boundary.) This problem has been treated before by various approaches [e.g., *Bremmer*, 1949; *Budden*, 1951; *Wait*, 1957; *Johler*, 1961]. In this paper we present computationally convenient formulas for the detailed calculation of the reflected wave excited by a given, arbitrarily polarized, plane wave incident obliquely upon a sharply bounded, magnetically anisotropic, absorptive ionosphere. A general method can be used for arbitrary emission direction with respect to the earth's magnetic field, but we have specialized certain formulas and the numerical results to the case of propagation perpendicular to the geomagnetic meridian.

Exact expressions (within the framework of the model described above) for the four possible reflection coefficients are readily obtained, and a computer program has been developed. The

relative simplicity of the method results from the use of impedance concepts coupled with *Booker's* [1938] treatment of the propagation vector. A vertically polarized incident plane wave has been assumed for the numerical calculations; hence only two of the four possible reflection coefficients have been computed. A wave frequency of 16 kc/sec has been assumed.

The following paragraphs outline the procedure followed to obtain expressions for the reflection coefficients, and present numerical results for the magnitude and phase of these coefficients as functions of the angle of incidence. Eight ionospheric models, which illustrate the effects of nonreciprocity and of variations in electron density and collision frequency have been considered.

PROCEDURE

Propagation parameters. It is assumed that an infinite plane electromagnetic wave propagating in free space is incident from below upon a plane, sharply bounded, anisotropic, absorptive ionosphere. The propagation geometry is illustrated in Figure 1. The quantities which ultimately describe the propagation characteristics of the resulting waves are as follows. Rationalized mks units are used.

e = electron charge (negative number).

m = electron mass.

N = number of electrons per cubic meter.

H_e = earth's magnetic field intensity.

¹Presently at the Physics Department, University of California at Los Angeles.

ν = electron collision frequency.
 k_0 = propagation constant in free space.
 $\omega = 2\pi f = 2\pi \times$ wave frequency.
 $\omega_H = \mu_0 \mathbf{H}_0 |e|/m$ = angular gyromagnetic frequency.
 $X = e^2 N / m \epsilon_0 \omega^2$.
 $\mathbf{Y} = \mu_0 \mathbf{H}_0 e / m \omega$.
 $Z = \nu / \omega$.
 $u = 1 - iZ$.
 θ = angle of incidence with respect to the vertical.
 $S = \sin \theta$, $C = \cos \theta$.
 ϕ = angle between the vertical and \mathbf{Y} .
 $c = (\epsilon_0 \mu_0)^{-1/2}$ = velocity of light in vacuum.
 η = characteristic admittance.
 $\eta_0 = \sqrt{\epsilon_0 / \mu_0}$ = characteristic admittance of free space.
 n = complex refractive index.
 Time convention: $e^{i\omega t}$.

We will first obtain various propagation parameters which enter into the reflection coefficient by following a treatment due to Booker

and a wave equation containing \mathbf{E} and \mathbf{P} . Substitution of equations 2 and 3 in the resulting wave equation leads to three equations in the (x_1, x_2, x_3) components and \mathbf{P} . A second, independent set of three equations in the components of \mathbf{P} are obtained by substitution of equations 2 into the Lorentz force equation (neglecting effect of the magnetic field of the wave in comparison with that of the earth):

$$-u\mathbf{P} = \epsilon_0 X \mathbf{E} + i\mathbf{P} \times \mathbf{Y}$$

The components of \mathbf{E} can be eliminated in these two sets of linear equations to obtain a set of three linear homogeneous equations in the components of \mathbf{P} . By equating to zero the determinant of this last set a quartic is obtained whose roots correspond to the four down-going waves of two magnetoionic components. In the special case of propagation perpendicular to the magnetic meridian, $Y_2 = 0$, the following expression is obtained

$$q^2 = C^2 - \frac{X(u-X)}{u(u-X) - \frac{1}{2}[Y_1^2 + S^2 Y_3^2] \pm [\frac{1}{4}(Y_1^2 + S^2 Y_3^2)^2 + Y_3^2(C^2 u - X)(u-X)]^{1/2}}$$

[1938]. In the ray approximation the wave function of a magnetoionic component for propagation in the x_2 - x_3 plane can be written as²

$$\exp [ik_0(ct - Sx_2 - qx_3)] \quad (1)$$

The only unknown in equation 1 is q , which is essentially the x_3 component of the propagation vector and which is a function of N , \mathbf{H}_0 , ω , θ , and ν . An expression for q in terms of these parameters can be obtained as follows: The polarization and electric vectors are expressed as

$$\mathbf{P} = \mathbf{A} \exp [ik_0(ct - Sx_2 - qx_3)] \quad (2)$$

For a given set of parameters, equation 1 has two values of q^2 , q_1^2 and q_2^2 , corresponding to the two transmitted magnetoionic components.

We now introduce two quantities, M and Q , which can be obtained from the linear equations in the components of \mathbf{P} and \mathbf{E} mentioned previously. M is the ratio of the x_3 and x_2 components of the polarization vector to the transmitted field of a magnetoionic component. Q is a similar ratio of the x_1 and x_2 components of the transmitted electric field. Again in the special case of propagation perpendicular to the magnetic meridian,

$$M_k = \frac{Y_3^2 [q_k^2 - C^2]^2 - [u(q_k^2 - C^2) + X][u(q_k^2 - C^2) + C^2 X]}{[u(q_k^2 - C^2) + X][iY_1(q_k^2 - C^2) - SXq_k]} \quad k = 1, 2$$

$$\mathbf{E} = \mathbf{B} \exp [ik_0(ct - Sx_2 - qx_3)] \quad (3)$$

The wave magnetic field vector, \mathbf{H} , is eliminated between the two Maxwell equations

$$\nabla \times \mathbf{E} = -\mu_0 \partial \mathbf{H} / \partial t \quad (4)$$

$$\nabla \times \mathbf{H} = \epsilon_0 \partial \mathbf{E} / \partial t + \partial \mathbf{P} / \partial t \quad (5)$$

² An alternative expression for (1) is

$$\exp \{ik_0[ct - (n \sin \theta')x_2 - (n \cos \theta')x_3]\}$$

where θ' refers to values of θ in the ionosphere. The form of (1) follows from this and from the Snell's law, $\sin \theta = n(\theta') \sin \theta'$, if q is defined as $q = n(\theta') \cos \theta'$.

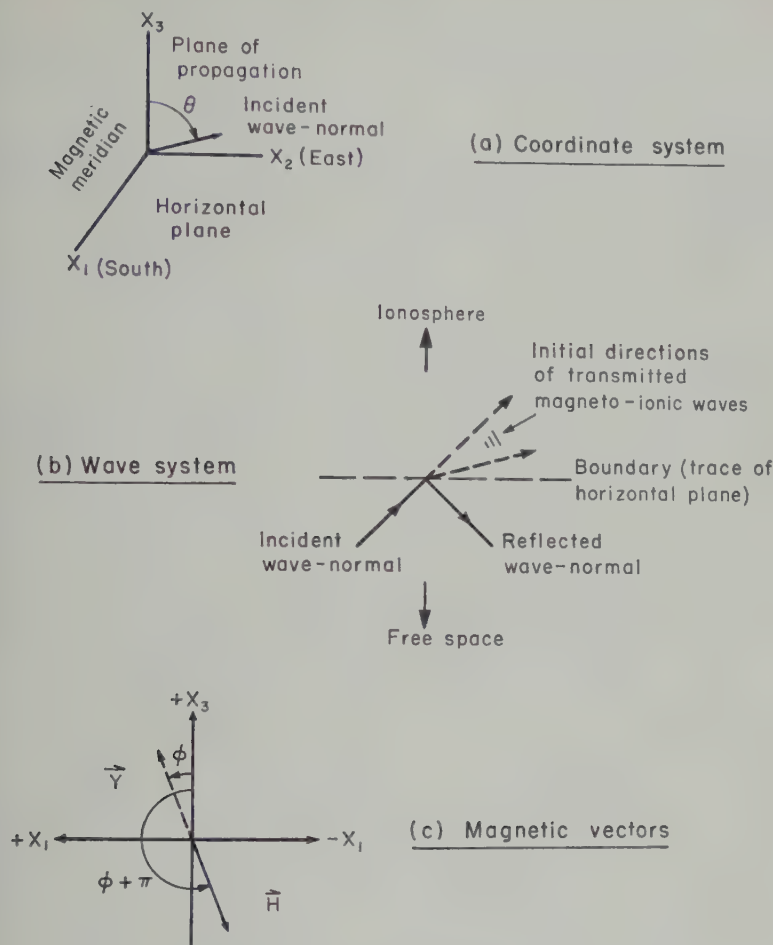


Fig. 1. Propagation diagrams.

$$Q_k = \frac{Y_3[q_k^2 - C^2]}{i[C^2 - Sq_k M_k][u(q_k^2 - C^2) + X]} \quad k = 1, 2 \quad (9)$$

and Q thus depend only upon the angle of incidence and the wave frequency for a given set of ionospheric parameters.³

Reflection coefficients. In order to obtain the

for the case of propagation at an arbitrary angle with the earth's field the quantities $M(q)$ and $Q(q)$ assume slightly more complicated forms than those of equations 8 and 9. The solution of the quartic equation for q in this case is much more complicated than for the case of $Y_2 = 0$. However, the subsequent quantities (from equation 10 on) are in the form presented.

for reflection coefficients we will use as boundary conditions the continuity of the tangential components of the total \mathbf{E} and \mathbf{H} fields. We introduce the following notation:

$(E_{\perp}, E_{\parallel}), (H_{\perp}, H_{\parallel})$ are components of the incident electric and magnetic field vectors perpendicular and parallel respectively to the plane of propagation.

$(E'_{\perp}, E'_{\parallel}), (H'_{\perp}, H'_{\parallel})$ are components of the reflected electric and magnetic field vectors perpendicular and parallel respectively to the plane of propagation.

$(E_{1,k}, E_{2,k}, E_{3,k})$ and $(H_{1,k}, H_{2,k}, H_{3,k})$, $k = 1, 2$, are the x_1, x_2 , and x_3 components of the electric and magnetic field vectors of the k th transmitted magnetoionic component.

It will be useful to introduce admittances into the boundary equations. Hence we first define appropriate admittances for the k th magnetoionic component as follows:

$$\eta_{\perp,k} = \frac{H_{2,k}}{E_{1,k}}; \eta_{\parallel,k} = -\frac{H_{1,k}}{E_{2,k}} \quad k = 1, 2 \quad (10)$$

In order to express these admittances in terms of ionospheric parameters, we substitute equation 3 into Maxwell's induction equation to obtain

$$\begin{aligned} (1/\eta_0)H_{1,k} &= -q_k E_{2,k} + S E_{3,k} \\ (1/\eta_0)H_{2,k} &= q_k E_{1,k} \quad k = 1, 2 \quad (11) \\ (1/\eta_0)H_{3,k} &= -S E_{1,k} \end{aligned}$$

Now, from equations 8, 10, 11, and the previously mentioned set of equations in the components of \mathbf{E} [cf. Booker, 1938], we find that

$$\begin{aligned} \eta_{\perp,k} &= \eta_0 q_k; \eta_{\parallel,k} \\ &= \eta_0 \left[q_k - S \left[\frac{-S q_k + [1 - q_k^2] M_k}{C^2 - S q_k M_k} \right] \right] \\ &\quad k = 1, 2 \quad (12) \end{aligned}$$

The boundary conditions expressed in terms of the above admittances are⁴

$$E_{\perp} + E'_{\perp} = E_{1,1} + E_{1,2} \quad (13)$$

$$E_{\perp} - E'_{\perp} = \frac{\eta_{\perp,1} E_{1,1} + \eta_{\perp,2} E_{1,2}}{C \eta_0} \quad (14)$$

$$E_{\parallel} - E'_{\parallel} = \frac{-[E_{2,1} + E_{2,2}]}{C} \quad (15)$$

$$E_{\parallel} + E'_{\parallel} = \frac{-[\eta_{\parallel,1} E_{2,1} + \eta_{\parallel,2} E_{2,2}]}{\eta_0} \quad (16)$$

To these we add the independent relations

$$Q_1 = E_{1,1}/E_{2,1} \quad Q_2 = E_{1,2}/E_{2,2} \quad (17)$$

If we assume that E_{\perp} and E_{\parallel} are known, all other field components may be determined from equations 13 to 17. The solution of these equations results in the following expressions for the

reflected wave:

$$E'_{\perp} = (\perp R_{\perp}) E_{\perp} + (\parallel R_{\perp}) E_{\parallel}$$

$$E'_{\parallel} = (\parallel R_{\parallel}) E_{\parallel} + (\perp R_{\parallel}) E_{\perp}$$

where the R 's are the reflection coefficients by

$$\parallel R_{\parallel} = \frac{\eta_0(a_1 - a_2) - C(\eta_{\parallel,2} a_1 - \eta_{\parallel,1} a_2)}{C \eta_0(a_1 b_2 - b_1 a_2)}$$

$$\perp R_{\perp} = \frac{b_2 Q_1 (C \eta_0 - \eta_{\perp,1}) - b_1 Q_2 (C \eta_0 - \eta_{\perp,2})}{C \eta_0(a_1 b_2 - b_1 a_2)}$$

$$\parallel R_{\perp} = \frac{a_1 Q_2 (C \eta_0 - \eta_{\perp,2}) - a_2 Q_1 (C \eta_0 - \eta_{\perp,1})}{C \eta_0(a_1 b_2 - b_1 a_2)}$$

$$\perp R_{\parallel} = \frac{\eta_0(b_2 - b_1) - C(b_2 \eta_{\parallel,1} - b_1 \eta_{\parallel,2})}{C \eta_0(a_1 b_2 - b_1 a_2)}$$

and where

$$a_k = Q_k + \frac{\eta_{\perp,k} Q_k}{C \eta_0};$$

$$b_k = -\left[1/C + \frac{\eta_{\parallel,k}}{\eta_0} \right] \quad k = 1, 2$$

It is of interest to note that as $H_0 \rightarrow \infty$, $q_1 \rightarrow q_2 = q$; $Q_1 \rightarrow Q_2 = Q$, etc.), equations 20 to 23 reduce to the following simpler

$$\parallel R_{\parallel} \rightarrow (C \eta_{\parallel} - \eta_0)/(C \eta_{\parallel} + \eta_0)$$

$$\perp R_{\perp} \rightarrow (C \eta_0 - \eta_{\perp})/(C \eta_0 + \eta_{\perp})$$

$$\parallel R_{\perp} \rightarrow \perp R_{\parallel} \rightarrow 0$$

These can be identified as the usual coefficients [Stratton, 1941] by using the fact that $S^2 + q^2 = n^2$.

Transmitted wave. The transmitted magnetoionic components at the boundary can be readily obtained from equations 13 to 17 and are given by

$$E_{1,1} = \frac{2Q_1(b_2 E_{\perp} - a_2 E_{\parallel})}{a_1 b_2 - a_2 b_1}$$

$$E_{1,2} = \frac{2Q_2(a_1 E_{\parallel} - b_1 E_{\perp})}{a_1 b_2 - a_2 b_1}$$

$$E_{2,1} = \frac{2(b_2 E_{\perp} - a_2 E_{\parallel})}{a_1 b_2 - a_2 b_1}$$

$$E_{2,2} = \frac{2(a_1 E_{\parallel} - b_1 E_{\perp})}{a_1 b_2 - a_2 b_1}$$

⁴ The 1, 2, 3 components of the incident and reflected electric fields expressed in terms of their perpendicular and parallel (\perp , \parallel) components are respectively: $(E_{\perp}, -E_{\parallel}C, E_{\parallel}S)$ and $(E'_{\perp}, E'_{\parallel}C, E'_{\parallel}S)$.

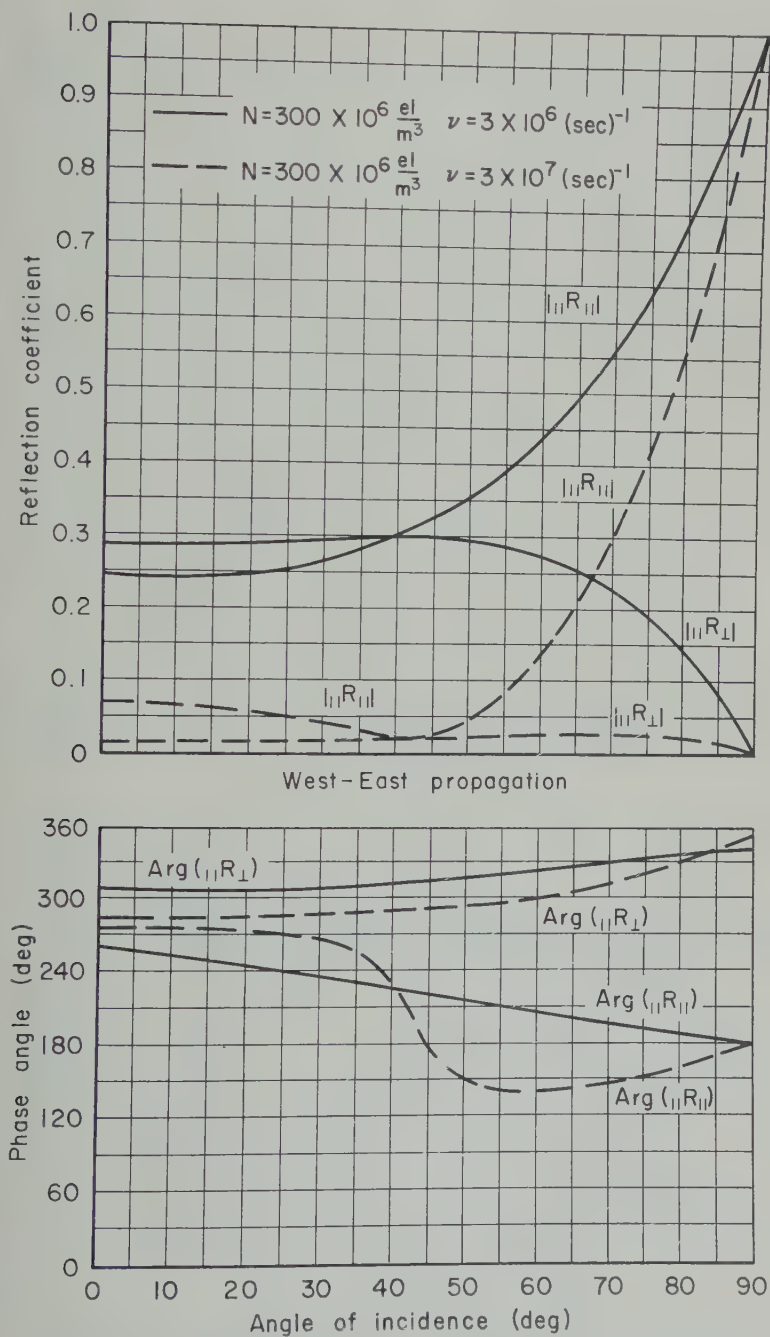


Fig. 2. Reflection coefficients for a high-density daytime ionosphere.

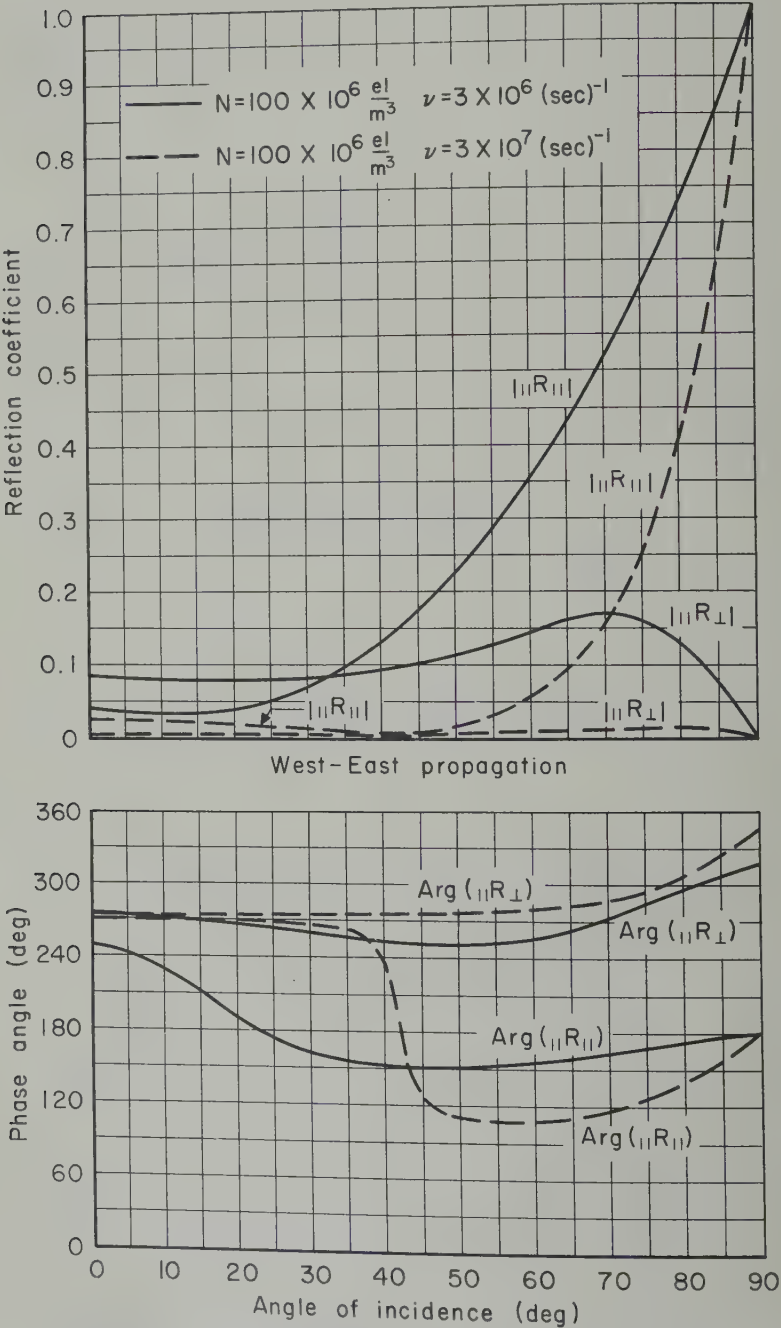


Fig. 3. Reflection coefficients for a low-density daytime ionosphere.

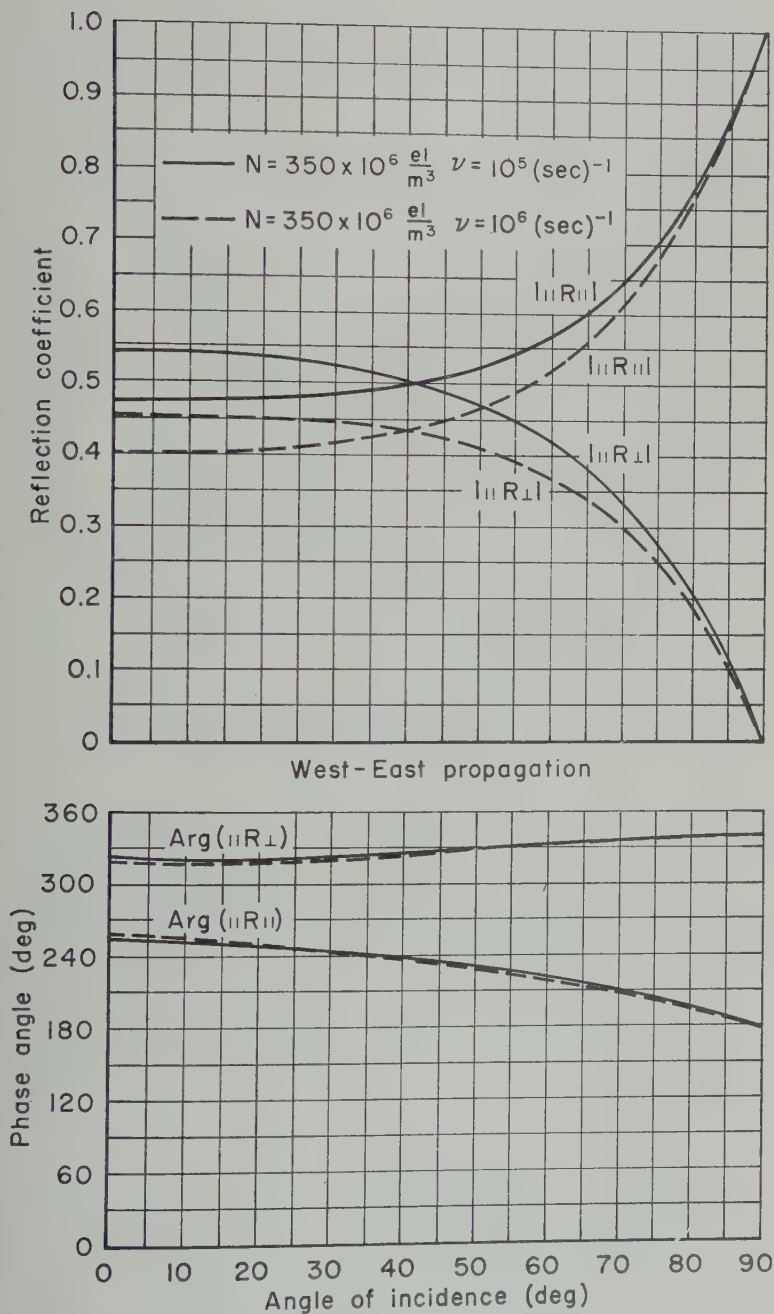


Fig. 4. Reflection coefficients for a high-density nighttime ionosphere.

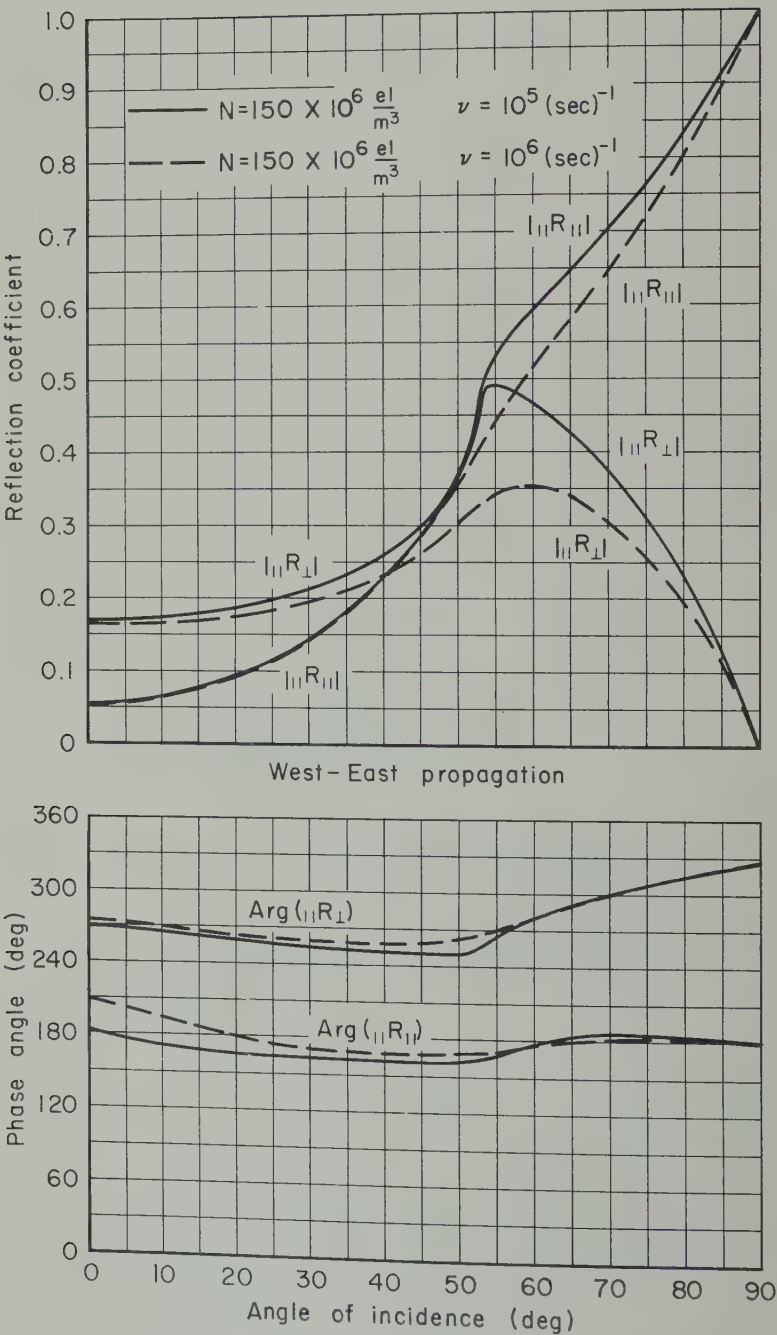


Fig. 5. Reflection coefficients for a low-density nighttime ionosphere.

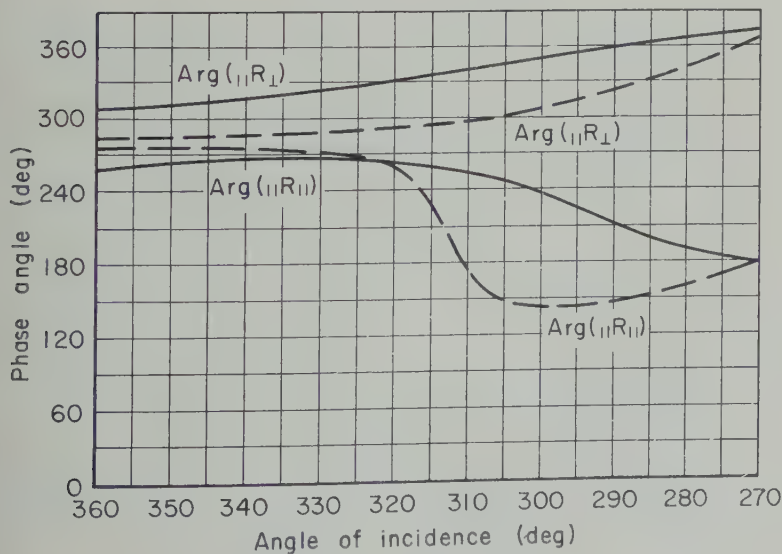
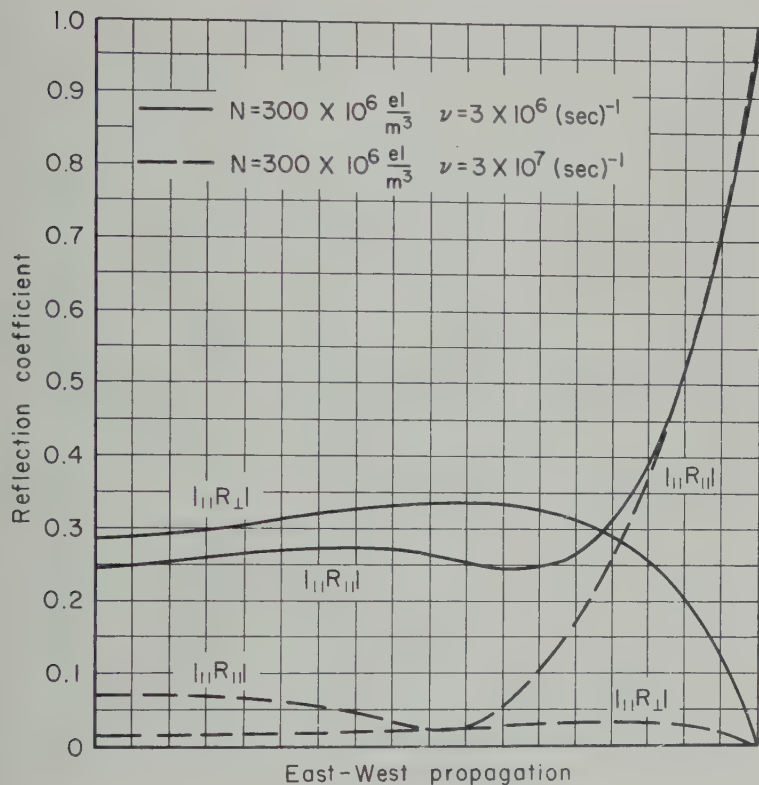


Fig. 6. Reflection coefficients for a high-density daytime ionosphere.

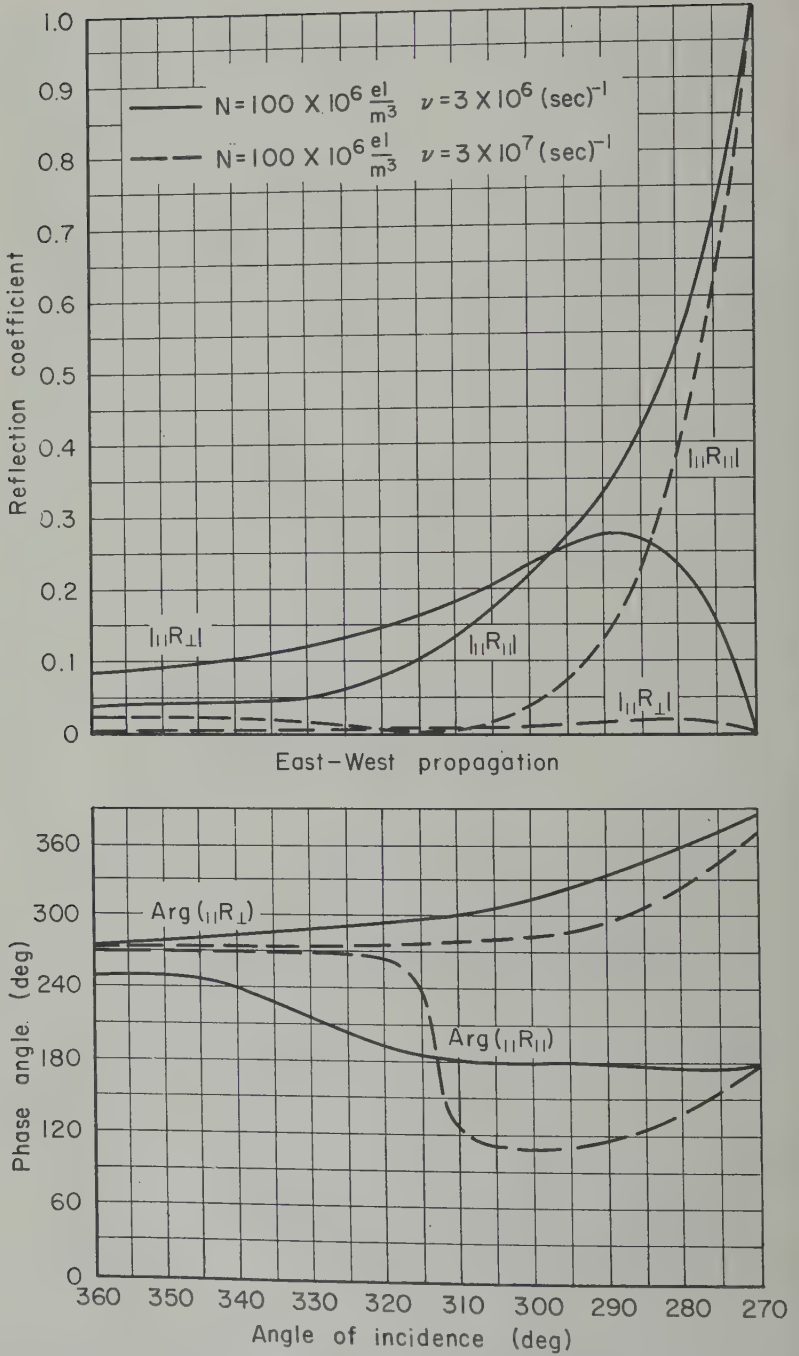


Fig. 7. Reflection coefficients for a low-density daytime ionosphere.

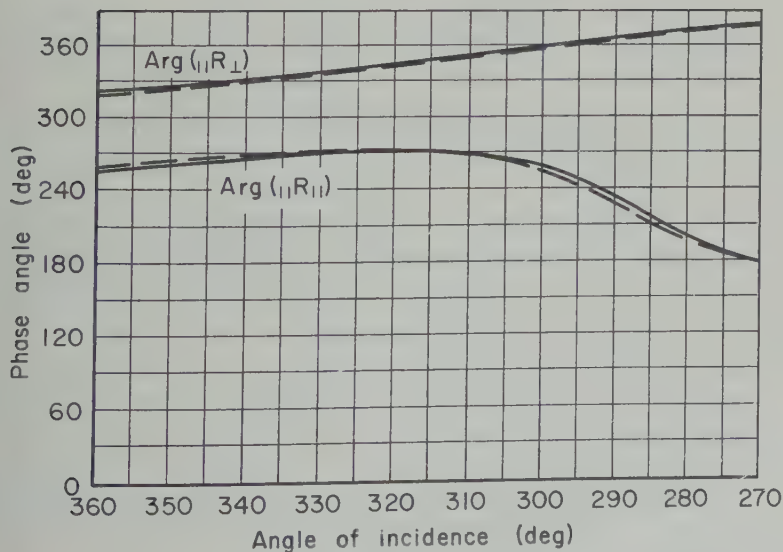
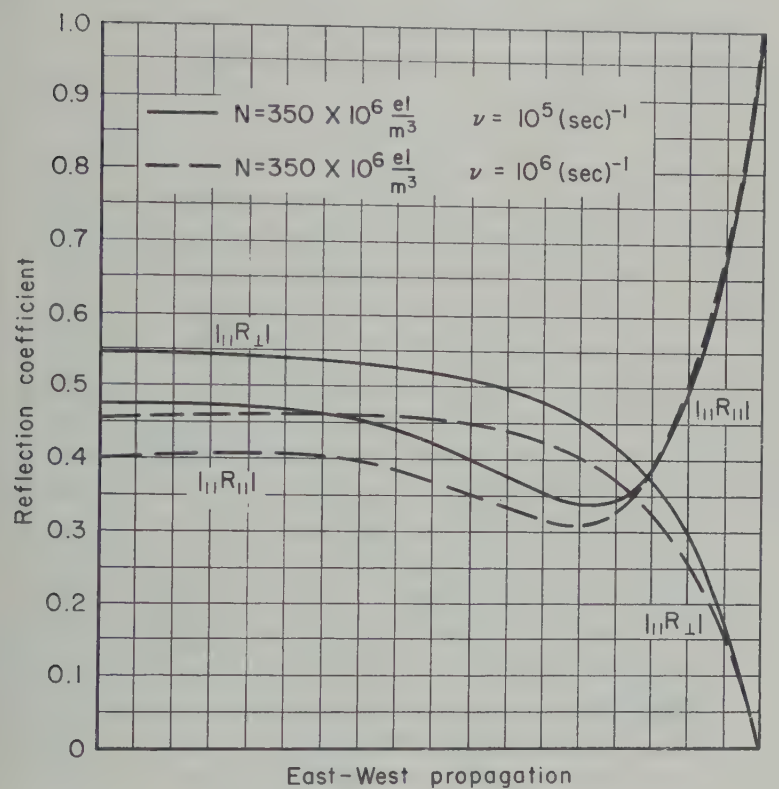


Fig. 8. Reflection coefficients for a high-density nighttime ionosphere.

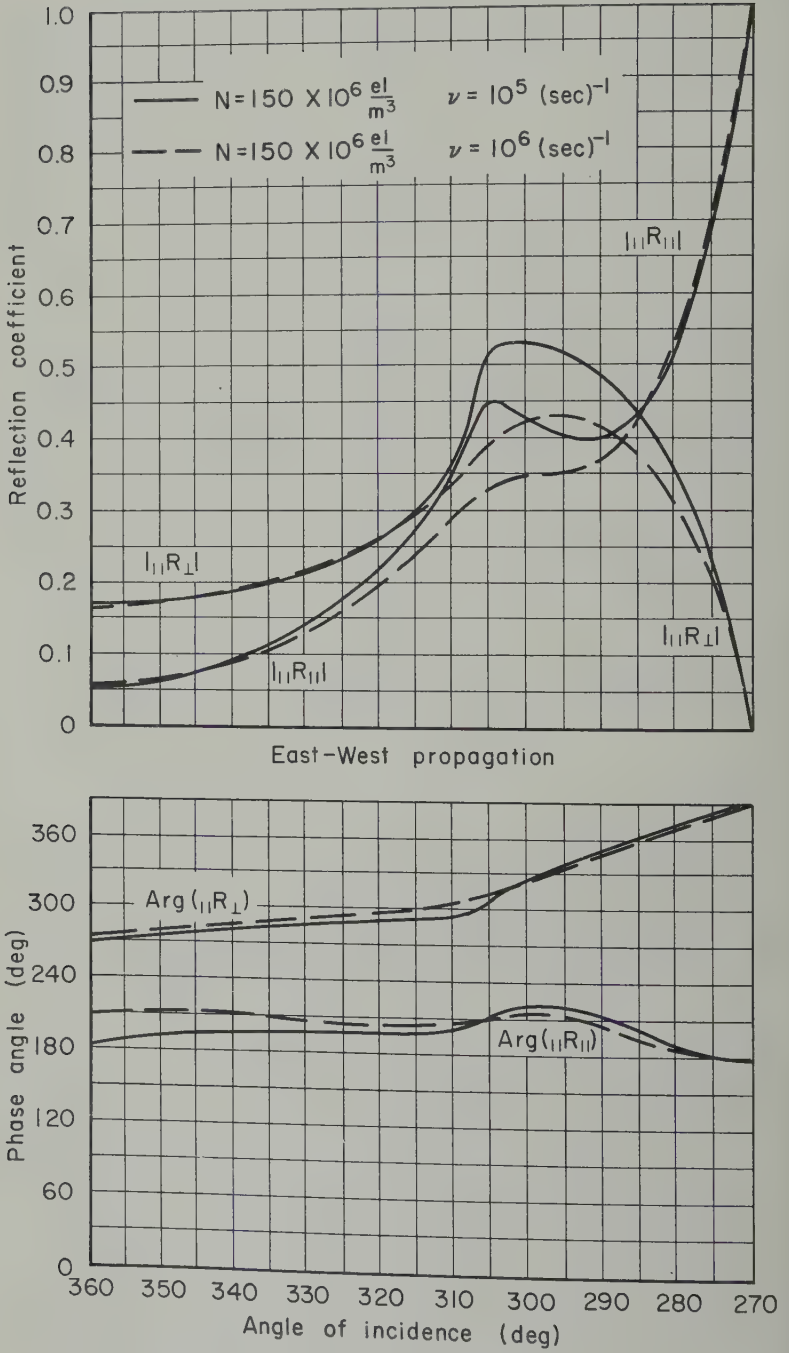


Fig. 9. Reflection coefficients for a low-density nighttime ionosphere.

TABLE 1. Ionospheric Models

incident 16 kc/sec E wave. Figures 2 to 5 are west \rightarrow east; Figures 6 to 9 are east \rightarrow west. $\omega_H = 7.85 \times 10^6$ rad./sec; propagation in northern hemisphere; $\phi = 20^\circ$.

	Daytime Model				Nighttime Model			
	Figs. 2, 6		Figs. 3, 7		Figs. 4, 8		Figs. 5, 9	
Electron density, el/m^3	$N = 300 \times 10^6$		$N = 100 \times 10^6$		$N = 350 \times 10^6$		$N = 150 \times 10^6$	
Collision frequency, $\nu \times 10^{-1}$	$\nu = 3 \times 10^6$	$\nu = 3 \times 10^7$	$\nu = 3 \times 10^6$	$\nu = 3 \times 10^7$	$\nu = 10^5$	$\nu = 10^6$	$\nu = 10^5$	$\nu = 10^6$
Boundary height, km	(solid)	(dashed)	(solid)	(dashed)	(solid)	(dashed)	(solid)	(dashed)
	70-71				88-90			

In these equations it can be seen that only the transmitted magnetoionic component is present if the polarization of the incident wave is such that at the boundary

$$E_{\perp}/E_{\parallel} = a_k/b_k \quad k = 1, 2$$

NUMERICAL RESULTS

In Figures 2 to 9 we present numerical results for our daytime and four nighttime ionospheric models. The magnitudes and phases of the reflection coefficients are plotted vs. angle of incidence. Each model is characterized by values of electron density and collision frequency (and thus boundary height). Figures 2 and 3 correspond to west \rightarrow east propagation, and Figures 6 to 9 correspond to east \rightarrow west propagation. The parameters are shown in Table 1. The values of electron density chosen are representative of upper and lower bounds to frequently used estimates [cf. *Waynick*, 1957; *Moler*, 1960]. For each electron density, two values of collision frequency are used to illustrate its effect on the reflection coefficients. We have followed *Phelps and Pack* [1959] as an approximate guide to the boundary height with collision frequency. This article presents two rather different procedures; in our figures the dotted lines are based on the older values given by them.⁵

DISCUSSION AND CONCLUSIONS

It is evident from equations 8, 9, and 12 that reciprocity exists between east \rightarrow west and

however, refer to *Phelps* [1960] for a further discussion of the validity of using 'effective' collision frequencies.

west \rightarrow east propagation, since S changes sign. This effect has been noted experimentally and theoretically [e.g. *Taylor*, 1960; *Johler*, 1961], and computations not reported here show very close agreement with results of *Johler* (allowing for different conventions in representing the incident and reflected waves). In the case, again, of propagation perpendicular to the magnetic meridian, the only differences in the reflection coefficients produced by changing from the northern to the southern hemisphere are 180° phase changes in $_{\parallel}R_{\perp}$ and $_{\perp}R_{\parallel}$. These phase changes result from the change of sign of Y_3 .

Reference to Figures 2 through 9 shows that in general, for the larger angles of incidence, the west \rightarrow east $_{\parallel}R_{\parallel}$ coefficients are larger than the east \rightarrow west ones. Although it is difficult to relate closely our values of reflection coefficients with experimentally observed long-distance attenuations, our results would seem to be compatible with the observations of *Taylor* [1960] that the attenuation is greater for westerly than for easterly propagation. There is an indication, however, that, at smaller angles of incidence, $_{\parallel}R_{\parallel}$ can become somewhat larger for east \rightarrow west than for west \rightarrow east propagation. For example, a comparison of Figures 2 and 6 substantiates this statement for angles of incidence between 0° and 30° .

We now relate our computations with experiments described by *Straker* [1955] and *Bain, Bracewell, Straker, and Westcott* [1952]. They report results at 16 kc/sec for a 90-km path between Rugby and Cambridge, England. This path is reasonably compatible with our west \rightarrow east, northern hemisphere, models. Their results, which correspond to angles of incidence near 30° ,

TABLE 2. Experimental Magnitudes of Reflection Coefficients

Reflection Coefficient	Daytime		Nighttime	
	Summer	Winter	Summer	Winter
$_{ }R_{ }$	0.15		0.50	
$_{ }R_{\perp}$	0.13	0.26	0.56	0.37

are summarized briefly for convenience in Table 2.

Interpretation of the experimental results quoted, in terms of the models used here, would indicate in every case that the higher, older, estimates of collision frequency result in low values for the reflection coefficients.* For the lower and newer collision frequencies, the computational results for the high-density daytime model are close to the experimental winter daytime results. Similar agreement exists between the computational results for the high-density nighttime model and the experimental summer nighttime results. In both low-density computational models, however, the agreement with experiments is not so close.

Acknowledgments. We are indebted to Dr. H. G. Booker for valuable advice at the early stages of this study. Our thanks are also due to E. M. Fairbrother for programming the numerical calculations.

REFERENCES

Bain, W. C., R. N. Bracewell, W. T. Straker, and B. A. Westcott, The ionospheric propagation of

* J. R. Wait (private communication) has suggested, however, that at highly oblique incidence, the daytime data indicates that ν is more like 10^7 .

radio waves of frequency 16 kc/s over distance of about 540 km, *Proc. IEE*, Monograph May 1952 [99 (C), p. 250], 1952.
Booker, H. G., Propagation of wave packet incident obliquely upon a stratified, doubly ionizing ionosphere, *Phil. Trans. Roy. Soc. A*, **A**, 237, 411, 1938.
Bremmer, H., *Terrestrial Radio Waves*, Interscience Publishing Co., Inc., New York, 1949.
Budden, K. G., The reflection of very low frequency radio waves at the surface of a bounded ionosphere with superimposed magnetic field, *Phil. Mag.*, **42**, 833, 1951.
Johler, J. R., Magneto-ionic propagation phenomena in low- and very-low-radio frequency reflected by the ionosphere, *J. Research NBS*, **65D**, 53, 1961.
Moler, W. F., VLF propagation effects of the ionospheric layer produced by cosmic rays, *Geophys. Research*, **65**, 1459-1468, 1960.
Phelps, A. V., Propagation constants for magnetic waves in weakly ionized, dry air, *Appl. Phys.*, **31**(10), 1723, 1960.
Phelps, A. V., and J. L. Pack, Electron collision frequencies in nitrogen and in the lower ionosphere, *Phys. Rev. Letters*, **3**, 340, 1959.
Straker, T. W., The ionospheric propagation of radio waves of frequency 16 kc/s over short distances, *Proc. IEE*, Monograph 114R, 1955 [102 (C), p. 122], 1955.
Stratton, J. A., *Electromagnetic Theory*, McGraw-Hill Book Company, New York, 1941.
Taylor, W. L., Daytime attenuation rates in the very low frequency band using atmospheric absorption, *J. Research NBS*, **64D**, 349, 1960.
Waynick, A. H., The present state of knowledge concerning the lower ionosphere, *Proc. IRE*, **45**, 741, 1957.
Yabroff, J. W., Reflection at a sharply bounded ionosphere, *Proc. IRE*, **45**, 750, 1957.

(Manuscript received March 29, 1961)

The Conductive 'Chapman' Layer

C. H. CUMMACK

Geophysical Observatory, Christchurch, New Zealand

Abstract. The shape of the electron production function, when ionizing radiation acts as an atmospheric heat source, is given for atmospheres in mixed or diffusive equilibrium.

Introduction. The theory of the absorption of monochromatic ionizing radiation in an isothermal atmosphere was stated by *Chapman* [1931] and extended by *Nicolet* [1951] to the case where the atmosphere has a linear gradient of temperature with height. The relations for an atmosphere with any temperature profile and the differential equation of atmospheric scale height when monochromatic radiation acts as an atmospheric heat source have been given by *Cummack* [1961].

It is the purpose of this paper to inquire more fully into the shape of the electron production function when the atmosphere is heated by the absorption of ionizing radiation and cooled by conduction downward.

Two cases will be considered: (a) the atmospheric gases are uniformly mixed; (b) the atmospheric constituents are in diffusive equilibrium. Radiative cooling has been considered important in this treatment.

Theory. The expression for the photon flux at any level in the atmosphere is [*Cummack*, 1961]

$$I = I_{\infty} \exp -(P\sigma \operatorname{ch} \chi) \quad (1)$$

where I_{∞} is the incident photon flux, P is the partial pressure of the ionizable component at that level, σ is related to the mass absorption coefficient, χ is the well-known 'Chapman' function. The total number of photons Q , absorbed at this level, per unit time, is thus

$$Q = I_{\infty}[1 - \exp -(P\sigma \operatorname{ch} \chi)] \quad (2)$$

It is now assumed that the heat input is directly proportional to the rate of electron production. This is not always strictly true, as it gives stationary values of electron density,

but it should suffice for a first approximation. Above any level defined by P , we may then write the heat gain, G , as

$$G = \gamma Q \quad (3)$$

where γ is the thermal energy released for the production and annihilation of each electron; and from equation 2

$$G = \gamma I_{\infty}[1 - \exp -(P\sigma \operatorname{ch} \chi)] \quad (4)$$

Now, if there is a conductive heat loss downward to a sink well below the maximum electron production, then the heat loss, L , is

$$L = K_0 \frac{dT}{dh} \quad (5)$$

where K_0 is the thermal conductivity, T is the atmospheric temperature, and h is vertical height.

Further, the thermal conductivity varies only as $T^{1/2}$, and equation 5 may be rewritten

$$L = K_1 T^{1/2} \frac{dT}{dh} \quad (6)$$

Under conditions of conductive equilibrium

$$L = G \quad (7)$$

therefore

$$K_1 T^{1/2} \frac{dT}{dh} = \gamma I_{\infty}[1 - \exp -(P\sigma \operatorname{ch} \chi)] \quad (8)$$

The solution of equation 8 gives the variation of atmospheric parameters with height, while the shape of the electron production function q is

$$q = q_0 \frac{H_0}{H} P \sigma \operatorname{ch} \chi \exp (1 - P\sigma \operatorname{ch} \chi) \quad (9)$$

where q_0 , H_0 are values of q , H at $P\sigma \operatorname{ch} \chi = 1$ [*Cummack*, 1961].

Solution for mixed atmosphere. For a mixed atmosphere, K_1 of equation 8 is a constant, while the atmospheric scale height H is the same for both the ionizable constituent and the atmosphere; thus without loss in generality T in equation 8 may be replaced by H . A mixed atmosphere under gravitational equilibrium satisfies the condition

$$dP/P = -dh/H \tag{10}$$

which may be integrated to give

$$P = P_0 \exp = \left(\int_0^h \frac{dh}{H} \right) \tag{11}$$

If we define the reduced height as

$$z = \int_0^h \frac{dh}{H} \tag{12}$$

then, from equation 11,

$$P\sigma \operatorname{ch} \chi = P_0\sigma \operatorname{ch} \chi \exp - z \tag{13}$$

Define the boundary conditions as

$$H = H_0 \quad \frac{dH}{dh} = \Gamma_0 \quad h = 0 \quad z = 0$$

at $P_0\sigma \operatorname{ch} \chi = 1$. Then by substituting equation 8 we get

$$\frac{1}{\Gamma_0} \left(\frac{H}{H_0} \right)^{1/2} \frac{dH}{dh} = \frac{1 - \exp - [\exp (-z)]}{1 - \exp (-1)}$$

Further, from equation 12

$$\frac{dH}{dh} = \frac{1}{H} \frac{dH}{dz}$$

and

$$\left(\frac{H_0}{H} \right)^{1/2} \frac{d}{dz} \left(\frac{H}{H_0} \right) = \Gamma_0 \frac{1 - \exp - [\exp (-z)]}{1 - \exp (-1)}$$

This may be integrated to give

$$\frac{H}{H_0} = [1 + \Gamma_0\phi(z)]^2$$

TABLE 1

h/H_m	$\Gamma_m = 0$	$\Gamma_m = 0.1$				$\Gamma_m = 0.2$				$\Gamma_m = 0.3$		
	q/q_m	q/q_m	H/H_m	Γ/Γ_m	Z	q/q_m	H/H_m	Γ/Γ_m	Z	q/q_m	H/H_m	Γ/Γ_m
∞			1.133				1.277				1.442	
4.0	0.049	0.067	1.128	0.040	3.662	0.086	1.267	0.052	3.352	0.108	1.415	0.064
3.5	0.079	0.102	1.126	0.061	3.218	0.126	1.261	0.075	2.976	0.153	1.403	0.090
3.0	0.128	0.157	1.122	0.094	2.773	0.185	1.251	0.111	2.578	0.211	1.387	0.127
2.5	0.207	0.236	1.116	0.145	2.327	0.265	1.238	0.163	2.176	0.294	1.364	0.182
2.0	0.321	0.351	1.107	0.221	1.877	0.346	1.218	0.239	1.769	0.405	1.331	0.257
1.5	0.484	0.508	1.094	0.334	1.423	0.530	1.189	0.348	1.354	0.549	1.285	0.363
1.0	0.693	0.704	1.073	0.497	0.961	0.712	1.146	0.505	0.926	0.722	1.220	0.512
0.9	0.737	0.744	1.068	0.535	0.868	0.750	1.136	0.542	0.839	0.760	1.204	0.549
0.8	0.780	0.783	1.062	0.578	0.774	0.791	1.125	0.583	0.750	0.797	1.187	0.587
0.7	0.818	0.826	1.056	0.623	0.680	0.830	1.112	0.626	0.661	0.833	1.168	0.629
0.6	0.862	0.865	1.050	0.670	0.585	0.866	1.099	0.672	0.570	0.868	1.149	0.673
0.5	0.900	0.900	1.043	0.720	0.489	0.900	1.086	0.720	0.479	0.902	1.128	0.720
0.4	0.932	0.931	1.035	0.772	0.393	0.932	1.071	0.771	0.386	0.930	1.106	0.770
0.3	0.959	0.959	1.027	0.826	0.296	0.956	1.055	0.824	0.292	0.958	1.082	0.822
0.2	0.981	0.978	1.019	0.883	0.198	0.974	1.038	0.880	0.196	0.980	1.056	0.877
0.1	0.992	0.995	1.010	0.940	0.100	0.994	1.020	0.939	0.099	0.994	1.029	0.936
0	1.000	1.000	1.000	1.000	0.000	1.000	1.000	1.000	0.000	1.000	1.000	1.000
-0.1	0.995	0.994	0.990	1.060	-0.101	0.993	0.979	1.062	-0.101	0.993	0.969	1.066
-0.2	0.979	0.977	0.979	1.121	-0.202	0.975	0.958	1.126	-0.204	0.973	0.936	1.134
-0.3	0.951	0.948	0.967	1.182	-0.305	0.943	0.934	1.192	-0.309	0.938	0.901	1.204
-0.4	0.912	0.904	0.955	1.241	-0.409	0.895	0.910	1.258	-0.418	0.886	0.864	1.278
-0.5	0.862	0.846	0.942	1.299	-0.514	0.830	0.884	1.324	-0.530	0.813	0.824	1.354
-0.6	0.801	0.776	0.929	1.355	-0.621	0.749	0.857	1.389	-0.645	0.719	0.782	1.430
-0.7	0.731	0.694	0.915	1.407	-0.730	0.654	0.829	1.452	-0.763	0.609	0.739	1.511
-0.8	0.653	0.604	0.901	1.456	-0.840	0.548	0.799	1.514	-0.887	0.484	0.692	1.594
-0.9	0.571	0.508	0.886	1.501	-0.952	0.436	0.768	1.572	-1.014	0.353	0.643	1.676
-1.0	0.488	0.412	0.871	1.539	-1.066	0.325	0.737	1.628	-1.147	0.229	0.591	1.769
-1.5	0.138	0.060	0.790	1.682	-1.668	0.011	0.559	1.914	-2.021	0.000	0.277	2.612
-2.0	0.012	0.001	0.704	1.787	-2.197	0.000	0.346	2.433	-3.036	—	—	—
			$h_s/H_m = -4.931$				$h_s/H_m = -2.681$				$h_s/H_m = -1.736$	
			$Z_s = -13.530$				$Z_s = -7.147$				$Z_s = -4.987$	

function, derived from the equation

$$= \frac{1}{2[1 - \exp(-1)]} \int_0^z \{1 - \exp - [\exp(-z)]\} dz \quad (18)$$

The function $\phi(z)$ is of some importance in this theory and occurs in modified form later. Also from equation 12

$$z = \int_0^h \frac{dh}{H}$$

$$h = \int_0^z H dz \quad (19)$$

$$\frac{h}{H_0} = \int_0^z [1 + \Gamma_0 \phi(z)]^2 dz \quad (20)$$

Equation 19 can be evaluated from tables of exponential integrals; equation 20 can be solved numerically. The associated electron production

$$\frac{q}{q_0} = \frac{H_0}{H} \exp(1 - z - \exp - z) \quad (21)$$

can now be calculated.

For most practical purposes q is better referred to the maximum of electron production than to the level $P\sigma$ at $\chi = 1$. Also h/H_m is a more natural measure than z .

Table 1 therefore gives the ratios of the required parameters to their values at the actual maximum. These have been interpolated from calculations using equations 17, 18, 20, and 21. Included in Table 1 are values of h_s/H_m and z_s referring to the lowest possible position of the heat sink. This occurs where $H/H_0 = 0$ and may easily be calculated from equations 17, 18, and 20.

Solution for diffusive equilibrium. When the atmosphere is in diffusive equilibrium, K_1 of equation 8 is not a constant. Following Johnson [1956], we may write

$$K_1 = K_2[3.5 + 6.7f] \quad (22)$$

where f is the fraction of atomic particles and

TABLE 1. Continued

H_m	$\Gamma_m = 0.4$				$\Gamma_m = 0.5$				$\Gamma_m = 0.6$			
	q/q_m	H/H_m	Γ/Γ_m	Z	q/q_m	H/H_m	Γ/Γ_m	Z	q/q_m	H/H_m	Γ/Γ_m	Z
∞		1.615				1.810				2.034		
4.0	0.131	1.571	0.077	2.907	0.155	1.735	0.092	2.719	0.178	1.906	0.106	2.555
3.5	0.177	1.553	0.106	2.587	0.203	1.709	0.122	2.428	0.228	1.870	0.137	2.290
3.0	0.239	1.528	0.144	2.262	0.266	1.673	0.162	2.133	0.291	1.823	0.178	2.019
2.5	0.321	1.494	0.198	1.932	0.348	1.627	0.214	1.830	0.372	1.762	0.231	1.741
2.0	0.429	1.451	0.212	1.592	0.454	1.564	0.289	1.517	0.475	1.683	0.302	1.451
1.5	0.567	1.382	0.375	1.239	0.584	1.480	0.385	1.198	0.597	1.580	0.397	1.148
1.0	0.733	1.293	0.518	0.866	0.742	1.367	0.525	0.838	0.749	1.440	0.531	0.813
0.9	0.767	1.272	0.553	0.788	0.774	1.339	0.560	0.764	0.781	1.407	0.563	0.743
0.8	0.802	1.249	0.591	0.708	0.806	1.311	0.595	0.689	0.813	1.373	0.597	0.671
0.7	0.836	1.225	0.631	0.627	0.841	1.280	0.634	0.611	0.844	1.336	0.635	0.597
0.6	0.871	1.199	0.673	0.545	0.872	1.247	0.675	0.532	0.875	1.296	0.675	0.521
0.5	0.903	1.171	0.719	0.460	0.904	1.212	0.720	0.451	0.907	1.255	0.718	0.443
0.4	0.932	1.141	0.769	0.373	0.934	1.175	0.767	0.367	0.935	1.211	0.765	0.362
0.3	0.959	1.109	0.820	0.284	0.959	1.136	0.819	0.280	0.960	1.163	0.816	0.278
0.2	0.980	1.075	0.876	0.193	0.980	1.093	0.875	0.191	0.980	1.112	0.872	0.190
0.1	0.994	1.039	0.936	0.098	0.994	1.048	0.935	0.098	0.994	1.058	0.933	0.097
0	1.000	1.000	1.000	0.000	1.000	1.000	1.000	0.000	1.000	1.000	1.000	0.000
-0.1	0.993	0.958	1.068	-0.102	0.993	0.948	1.071	-1.103	0.992	0.938	1.074	-0.103
-0.2	0.972	0.914	1.141	-0.209	0.971	0.893	1.149	-0.212	0.970	0.871	1.158	-0.214
-0.3	0.934	0.867	1.219	-0.321	0.930	0.833	1.235	-0.328	0.928	0.799	1.251	-0.333
-0.4	0.876	0.817	1.301	-0.440	0.865	0.769	1.329	-0.453	0.855	0.720	1.361	-0.466
-0.5	0.795	0.763	1.390	-0.566	0.773	0.700	1.435	-0.588	0.751	0.635	1.490	-0.613
-0.6	0.686	0.706	1.486	-0.703	0.647	0.625	1.558	-0.740	0.604	0.541	1.652	-0.783
-0.7	0.556	0.644	1.591	-0.851	0.492	0.544	1.703	-0.910	0.416	0.436	1.872	-0.988
-0.8	0.407	0.578	1.709	-1.015	0.315	0.454	1.890	-1.114	0.203	0.314	2.228	-1.255
-0.9	0.256	0.507	1.846	-1.199	0.145	0.353	2.160	-1.359	0.030	0.160	3.132	-1.685
-1.0	0.125	0.430	2.018	-1.413	0.028	0.235	2.655	-1.700	—	—	—	—
-1.5	—	—	—	—	—	—	—	—	—	—	—	—
-2.0	—	—	—	—	—	—	—	—	—	—	—	—
	$h_s/H_m = -1.354$ $Z_s = -3.869$				$h_s/H_m = -1.119$ $Z_s = -3.208$				$h_s/H_m = -0.962$ $Z_s = -2.734$			

is a function of height. With any chosen atmosphere f may be expressed in terms of z , thus modifying only $\phi(z)$ of equation 17.

As the parameters of atmospheric composition are not well known, there seems little point in solving these equations over a range of conditions; only indications of the possible effects of a modified conductivity expression will be given.

Consider an atmosphere consisting of an ionizable component of atomic oxygen with scale height H , together with molecular nitrogen having a scale height H/k .

The height variation of oxygen may be expressed by

$$P_{[O]} = P_{[O],0} \exp - z \tag{23}$$

and that of nitrogen by

$$P_{[N_2]} = P_{[N_2],0} \exp - kz \tag{24}$$

and at any height Z

$$f = \frac{P_{[O]}}{P_{[O]} + P_{[N_2]}} \tag{25}$$

$$= \frac{1}{1 + a \exp - (k - 1)z} \tag{26}$$

where

$$a = P_{[N_2],0}/P_{[O],0}$$

Since T is defined everywhere by H , equation 8 then becomes

$$\frac{d}{dz} \left(\frac{H}{H_0} \right)^{1/2} = \frac{\Gamma_0}{2} \frac{1 - \exp - [\exp(-z)]}{1 - \exp(-1)} \cdot \frac{3.5 + \frac{6.7}{1+a}}{3.5 + \frac{6.7}{1+a \exp - (k-1)z}}$$

The solution with

$$\phi(z) = \frac{1}{2} \int_0^z \frac{1 - \exp - [\exp(-z)]}{1 - \exp(-1)} \cdot \frac{3.5 + \frac{6.7}{1+a}}{3.5 + \frac{6.7}{1+a \exp - (k-1)z}} dz$$

is now the same as before.

$\phi(z)$ is tabulated in Table 2 (for various values of a). This, together with equations 20, and 21, is sufficient for the calculation of the atmospheric parameters and the q function.

It is obvious that the shape of the q function given in Table 1 will not be greatly affected by a refinement in the conductivity expression, although the variations of atmospheric parameters could be considerably different for large negative values of z .

TABLE 2

$\phi(Z)$						
Z	Mixed Atmosphere	$a = 0.1$	$a = 0.2$	$a = 0.4$	$a = 0.8$	$a = \infty$
4.0	0.616	0.600	0.586	0.566	0.543	0.521
3.5	0.607	0.591	0.578	0.558	0.536	0.514
3.0	0.591	0.577	0.564	0.546	0.525	0.503
2.5	0.567	0.553	0.542	0.525	0.506	0.484
2.0	0.527	0.515	0.506	0.491	0.475	0.458
1.5	0.463	0.454	0.447	0.436	0.424	0.412
1.0	0.364	0.359	0.354	0.347	0.341	0.335
0.5	0.214	0.213	0.211	0.209	0.207	0.205
0.0	0.000	0.000	0.000	0.000	0.000	0.000
-0.5	-0.286	-0.289	-0.292	-0.296	-0.300	-0.304
-1.0	-0.632	-0.651	-0.665	-0.683	-0.699	-0.715
-1.5	-1.015	-1.066	-1.103	-1.148	—	—
-2.0	-1.409	-1.517	-1.590	-1.675	—	—
-2.5	-1.804	-2.001	-2.125	—	—	—
-3.0	-2.200	-2.524	-2.712	—	—	—
Diffusive level	—	-3.994	-3.070	-2.146	-1.221	-0.300

These tables are intended to be used as the basis for more accurate analysis of the transition region when deriving atmospheric parameters.

Acknowledgment. This work is part of the research program of the Christchurch Geophysical Observatory of the New Zealand Department of Scientific and Industrial Research, Superintendent, J. W. Beagley. Thanks are due to Miss C. L. Fryman and Miss E. Y. Currie for assistance in computation of Table 1, and to Mr. G. A. M. King for many helpful discussions.

REFERENCES

Chapman, S., The absorption and dissociative or

ionizing effect of monochromatic radiation in an atmosphere on a rotating earth, *Proc. Phys. Soc. London*, **43**, 26-45, 1931.

Cummack, C. H., Extensions of the 'Chapman' theory of layer formation, *J. Geophys. Research*, **66**, 1685-1697, 1961.

Johnson, F. S., Temperature distribution of the ionosphere under control of thermal conduction, *J. Geophys. Research*, **61**, 71-76, 1956.

Nicolet, M., Effects of the atmospheric scale height gradient on the variation of ionization and short wave absorption, *J. Atmospheric and Terrestrial Phys.*, **1**, 141-146, 1951.

(Manuscript received May 19, 1961;
revised June 9, 1961.)



Analysis of the F_1 - F_2 Transition Region

G. A. M. KING

*Department of Scientific and Industrial Research, Geophysical Observatory
Christchurch, New Zealand*

Abstract. The paper describes a method of deducing rates of electron production and loss and the atmospheric scale heights at the height of the transition between the F_1 and F_2 layers of the ionosphere. The electron density profile obtained from the ionogram is transformed into dimensionless coordinates and compared with a set of theoretical curves parametric in a factor, G , which determines the shape of the transition region. The assumptions under which the theoretical curves have been derived are clearly set out. The method is accurate enough to permit the study of changes in the ionosphere during a day.

Introduction. From ionospheric records the electron distribution in height can be obtained. It can be compared with the distribution expected theoretically from consideration of the processes of electron production and loss, and the comparison will yield values for the basic ionospheric parameters, namely, the rates of production and loss and atmospheric scale heights.

Previous studies of this kind on the F_1 - F_2 transition region were by Yonezawa, Takahashi, and Hirsh [1959] and by Hirsh [1959]. They used ionospheric electron density data and made some deductions about the average rates of electron production and loss.

This paper describes a method by which very sparse electron density data from individual ionograms can be treated to give the ionospheric parameters in the F region with an accuracy sufficient for the detailed analysis of sequences of ionograms. The method was developed for the study of daytime ionospheric storms but is also applicable for diurnal, seasonal, and sunspot cycle studies.

Necessary

height, kilometers.

scale height of the atomic oxygen in the ionosphere.

gradient of scale height, dH/dh .

'reduced' height, $\int dh/H$.

$\epsilon = e^{-Z}$.

electron density.

electron production function.

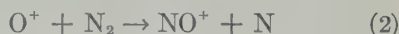
electron loss function.

electron loss coefficient in the 'quadratic' region; it is the rate coefficient of the

equation



β = electron loss coefficient in the 'linear' region; it is given by the product of the number density of molecular nitrogen and the rate coefficient of the equation



$r = \beta/\alpha n$.

$G = \beta_0^2/q_0\alpha$, where β_0 and q_0 are the values of β and q at $Z = 0$.

k = ratio of the molecular weight of N_2 to the atomic weight of O.

$k' = (k + \Gamma)/(1 + \Gamma)$.

f_n = plasma frequency.

$D = dh/(d \log f_n)$.

$\lambda = (d \ln n)/(dZ)$.

The reference level for Z , $Z = 0$, is defined here as the height of the maximum in the q function. This choice differs from that commonly made (maximum in the q function for overhead sun), but leads to simpler equations which do not involve the solar zenith angle.

Natural logarithms are written \ln ; logarithms to the base 10 are written \log . The conversion factor is

$$M = \log e = 0.4343 \quad (3)$$

Basic theory. In order that they may be inspected critically, the fundamental assumptions are listed:

1. The atmosphere is in diffusive equilibrium under gravity over the whole F_1 - F_2 transition region.

2. The gradient of scale height, Γ , is constant. This assumption is made only to simplify the numerical calculations.

3. The value of Γ is deduced from a conductive model of the atmosphere.

4. F -region ionization is produced by the absorption of monochromatic solar radiation.

5. The solar radiation is absorbed by atomic oxygen.

6. Electron loss occurs through dissociative recombination of nitric oxide ions (see equation 1 in Glossary).

7. The nitric oxide ions are formed by atom-ion interchange between the ionized atomic oxygen and nitrogen molecules (see equation 2 in Glossary).

8. The electron density is in equilibrium between production and loss.

9. The 'quadratic' loss coefficient, α , has the constant value 10^{-8} sec⁻¹/ion.

Since the particular loss processes assumed here, involving molecular nitrogen, seem to be much more important than the similar processes involving molecular oxygen [King and Roach, 1961], the latter have not been included in the analysis.

The detailed derivation of the various expressions from these assumptions is lengthy and will not be reproduced; the expressions themselves will simply be quoted.

The production function is [Nicolet, 1951]

$$q = q_0 \exp [(1 + \Gamma)(1 - Z - \epsilon^{-Z})] \quad (4)$$

The loss function is [Ratcliffe, 1956]

$$L = \beta_n / (1 + \Gamma) \quad (5)$$

where β depends on the number density of molecular nitrogen, proportional to its partial pressure and inversely proportional to its scale height. From the assumption of diffusive equilibrium between N_2 and O, it follows that

$$\beta = \beta_0 \exp [-(k + \Gamma)Z] \quad (6)$$

The equilibrium assumption allows equations 4 and 5 to be equated and solved for $n(Z)$, so that comparison can be made with experimental data. To make full use of the most accurate experimental data, however, it is preferable to use, instead of n ,

$$\lambda = \frac{d \ln n}{dZ}$$

This is most readily obtained by $(d \ln q)/dZ$ from (4) and $(d \ln L)/dZ$ from (5), eliminating $(d \ln \beta)/dZ$ with the help of (6). The result is

$$\lambda = (1 + \Gamma) \frac{k' - x(1 + r)}{1 + 2r}$$

For numerical computation of $\lambda(Z)$, it is to get r from its definition, $\beta/\alpha n$, and eqn. (5), replacing L by q from (4), in conformity with the equilibrium assumption, and inserting β from (6). This gives

$$r(r + 1)$$

$$= G \exp [-(1 + \Gamma)\{(2k' - 1)Z + 1\}]$$

Notice that the effect of the scale height is to reduce the effective mass ratio of atmospheric constituents from k to k'

$$k' = \frac{k + \Gamma}{1 + \Gamma}$$

The function $\lambda(Z)$ is plotted in Figure 1 for the case where $\Gamma = 0.3$. (Tabulations from these graphs [for $\Gamma = 0, 0.2, 0.3, 0.4$], which can be prepared are available on request.) The line in the figure represents an electron density profile about the F_1 - F_2 transition region for different values of G . It is interesting to note that, in these coordinates, there is no difference between a winter-time density profile at sunspot maximum ($\log G \doteq 1$) and a summer-time profile at sunspot minimum ($\log G \doteq 2$); of course, the corresponding $h'(f)$ look vastly different. $\log 0$ is representative of severe storm conditions when the F_2 layer is unobservable because $f_o F_2 < 3$ while the line labeled ∞ corresponds to the 'quadratic' region.

Method of analysis. Now, consider how the experimental data to these graphs can be obtained by the method $n(h)$ analysis used [King, 1954]

$$D(h) = \frac{dh}{d \log f_n}$$

This is related to λ by

$$\lambda = \frac{2H}{M D}$$

Also, h is related to Z by

$$Z = \int \frac{dh}{H}$$

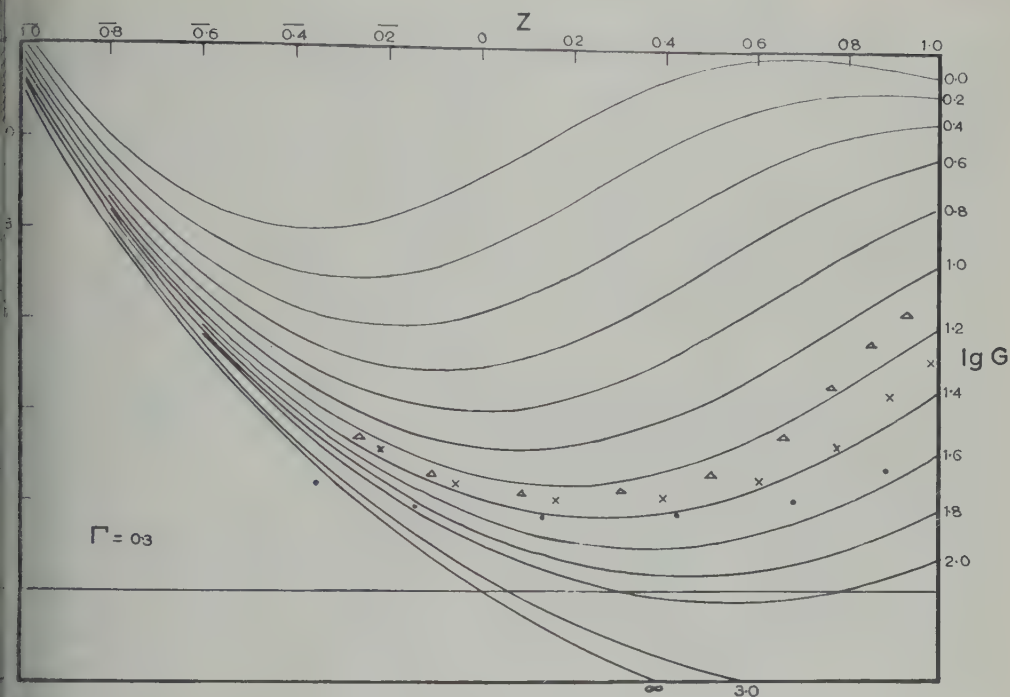


Fig. 1. The function $\lambda(Z)$ for $\Gamma = 0.3$. The points plotted on the figure are discussed in the text.

procedure, then, is as follows:

Choose a value of Γ (see later).

Set up a scale of $H(h)$ compatible with this of Γ .

Using this H scale convert the experimental n into values of Z (equation 11), adopting arbitrary zero.

Convert the experimental values of D into values of λ (equation 10).

From steps 3 and 4, plot $\lambda(Z)$ on a transparent sheet to the same scale as the theoretical λ (based, of course, on the adopted Γ).

Slide the transparent sheet over the theoretical graph, maintaining the λ scale but the Z scale (for Z has an arbitrary zero), the position of best fit is found.

If the fit is not good, repeat the whole process with a new scale of $H(h)$.

The method is illustrated by the points plotted in Figure 1; they represent a portion of the $n(h)$ measured about the F_1 - F_2 transition region at March 9, 1955, at Christchurch. The first trial, based on a value of $H = 30$ km at a height of 30 km (the adopted arbitrary zero in Z), is shown by the dots; it is not a good fit. The second

trial, $H_{177} = 40$ km (triangles), is much better; H seems a little too high. The third trial, $H_{177} = 37$ km (crosses), is better than the second at the higher values of Z but perhaps not quite as good at the low values of Z .

Before considering the reasons for the difficulty in getting a perfect fit (they are related to the choice of Γ) it is instructive to set out the results of this experiment in tabular form. See Table 1.

The values of β_0 and q_0 were deduced from G by

$$G = r_0(r_0 + 1) \quad (12)$$

$$= \beta_0^2 / q_0 \alpha \quad (13)$$

and

$$r_0 = \beta_0 / \alpha n_0 \quad (14)$$

n_0 being obtained by interpolation of the experimental data.

The product $q_0 H_0$ is of particular interest, as it is proportional to the flux of ionizing radiation from the sun.

Choice of Γ . In theory, the fit of the experimental data to the graphs could determine which graph should be used, and hence the value of Γ . However, experimental error readily

TABLE 1

H_{177}	30 km	37 km	40 km
Z_{177}		+0.40	+0.30
True zero of Z	not satisfactory		
H_0		163 km	166 km
$\log G$		33 km	37 km
β_0		1.3	1.25
q_0		$8.3 \times 10^{-3}/\text{sec}$	$8.0 \times 10^{-3}/\text{sec}$
$q_0 H_0$		$344/\text{sec}$	$355/\text{sec}$
		$1.14 \times 10^9/\text{cm}^2$	$1.30 \times 10^9/\text{cm}^2$
		sec	sec

masks the small differences by which such a judgment might be made. Now, it seems reasonable to assume that the high values of H in the F region are maintained against loss of heat by conduction to the lower atmosphere; the heat sink can be taken to be near 105 km, where $H \doteq 10$ km. As the heat flow depends on Γ and on the thermal conductivity, which in turn is proportional to \sqrt{H} , a constant heat flow would mean that

$$\begin{aligned} \Gamma \sqrt{H} &= \sqrt{H} \frac{dH}{dh} \\ &= C \end{aligned} \tag{15}$$

Integrating,

$$H^{3/2} = \frac{3}{2}ch + \text{integration constant} \tag{16}$$

It is convenient, instead of adopting $H_{105} = 10$ km, to take as reference $H_{100} = 0$; in practice the difference is trivial. Equations 15 and 16 can then be solved to give Γ .

$$\Gamma = \frac{2H}{3(h - 100)} \tag{17}$$

Thus, if H is known at any height, Γ can be calculated. In the example illustrated in Figure 1, if $H_{177} = 37$ km, then Γ_{177} is 0.32, justifying use of the $\Gamma = 0.3$ graph.

Discussion of difficulties. It is clear that the fit of the data presented on Figure 1 cannot be improved by another choice of H ; the basic assumptions must be examined.

One of the most restrictive assumptions is that the gradient of scale height, Γ , is constant. In fact, even if the heat flow were constant, equation 15 requires that Γ decrease with increasing height, although the change is not large over the

height range considered. Now, a constant flow would obtain if the energy were supplied from outside the ionospheric atmosphere, unlikely, however, that the energy is supplied the same solar radiation that generates ionization in the F region. The heat flow decreases, therefore,

$$\Gamma \sqrt{H} \propto \int_h^\infty q \, dh$$

and, as the integral decreases rapidly with height above the maximum in the q function, will fall. The record used in Figure 1 has been analyzed more thoroughly with a variable height gradient given by equation 18, and the fit to the theoretical curves was much improved. The results of this very tedious calculation are given in Table 2.

The accuracy is such that it is possible to distinguish between $H_{177} = 38.5$ and $H_{177} = 37$ km. Notice that Γ_{177} is less than that deduced from equation 17, because, of course, not all the radiation has been absorbed at this height.

Another minor source of error is apparent in this more precise experiment; it is the effect of the nonlinearity of D in the $n(h)$ reduction, and subsequent interpolation to obtain D at the mean values of D over a height range in the $n(h)$ reduction. The nonlinearity shows a very slight flattening at the lowest values where the curvature is greatest.

In order to find other difficulties in the reduction, consider the example in Figure 2; the data are from a portion of the $n(h)$ profile above the F_1 - F_2 transition region at 1330, October 1949, at Christchurch. The first trial, based on $H_{206} = 80$ km, has not been included in this figure, being a very poor fit. The second trial, $H_{206} = 65$ km (dots), gives a reasonable fit to part of the data but shows somewhat irregular behavior about $Z = 0$, and is quite

TABLE 2

H_{177}	38.5 km	Γ_{177}	0.28
Z_{177}	0.38	$\log G$	1.32
True zero of Z			
r_0	163 km	r_0	4.1
n_0	$2.09 \times 10^5/\text{cc}$	β_0	8.6×10^{-3}
H_0	33.8 km	q_0	$351/\text{cc}$
$q_0 H_0$	$1.19 \times 10^9/\text{cm}^2 \text{ sec}$		

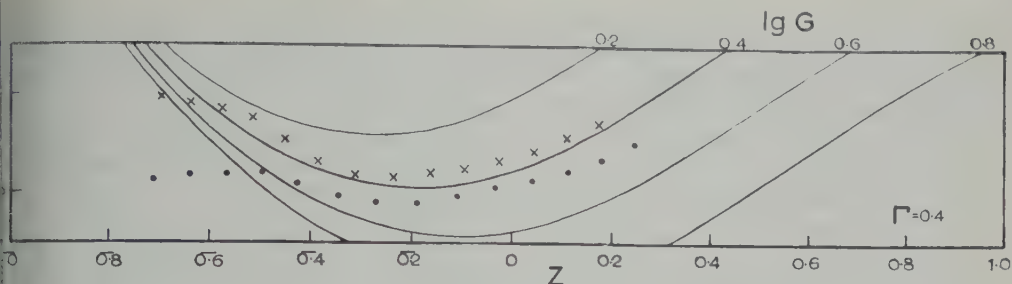


Fig. 2. Analysis of the $n(h)$ profile, 1330, October 14, 1959, at Christchurch.

very below $Z = -0.5$; although some improvement is possible with a slightly larger value of Γ , these defects will remain. The cause of the irregularities near $Z = 0$ is apparent when the original ionogram is re-examined—there is overlap of the ordinary and extraordinary traces in the region and above, so that the virtual heights in the $n(h)$ analysis could not be read accurately enough. The data were reread with a height correction (up to 3 km in h') for the overlap, and subsequent tests use these corrected

two explanations offer themselves for the defect below $Z = -0.5$: (1) there is another source of ionization besides that assumed for the region; (2) there may be a valley between the ordinary F_1 layers which has not been allowed for in the $n(h)$ reduction (normal reduction gives an equivalent monotonic electron distribution).

In the present state of knowledge, all that can be said of the first explanation is that it is possible. For example, if the bulk of the F region is due to the radiation of He II at 304 Å, there might be a contribution at these lower heights to the He I radiation at 584 Å; perhaps the ionization continuum also contributes.

The second explanation, that there may be a valley between the E and F_1 layers, can be tested immediately. The third trial with the profile illustrated in Figure 2 (crosses) uses an $n(h)$ profile which includes a shallow valley 34 km wide. The arbitrary reference for Z was chosen at the same value of electron density, and the height was raised by inclusion of the valley 34 km; the trial is based on $H_{215} = 65$ km. The general shape of the profile shows a marked improvement over the second trial which does not have a valley. In fact, by a somewhat wider valley still, one that has some depth and is

with the theoretical shape over the whole of the range considered. The question now arises how determined an effort should be made to secure agreement. The author believes it unwise to pursue the question too closely at present, preferring to admit that either of the explanations offered is possible.

It is interesting, though, to tabulate the results of the second and third trials shown in Figure 2. See Table 3.

Even a 34-km valley has not made a large percentage difference in the values obtained; the greatest change is in the height of maximum electron production, which has been raised 13 km by including the valley.

Conclusions. Comparison of Tables 2 and 3 points out some important facts (as the solar zenith angles for the two times illustrated are similar, a direct comparison is possible):

1. The scale height at the maximum of electron production is greater at sunspot maximum than at sunspot minimum (a factor of 2 for these particular records).

2. The maximum electron production is greater at sunspot maximum (a factor of 4).

TABLE 3

	No Valley, Uncorrected for Overlap	34-km Valley, Corrected
True zero of Z	218 km	231 km
H_0	70 km	71 km
n_0	$5.06 \times 10^5/\text{cc}$	$5.25 \times 10^5/\text{cc}$
$\log G$	0.45	0.37
r_0	1.25	1.11
β_0	$6.3 \times 10^{-3}/\text{sec}$	$5.8 \times 10^{-3}/\text{sec}$
q_0	$1.41 \times 10^3/\text{cc sec}$	$1.44 \times 10^3/\text{cc sec}$
$q_0 H_0$	$0.99 \times 10^{10}/\text{cm}^2$ sec	$1.02 \times 10^{10}/\text{cm}^2$ sec

3. The product, $q_0 H_0$, which is a measure of the radiation from the sun, is very much greater at sunspot maximum (a factor of 8).

4. The height of maximum electron production is greater at sunspot maximum; this is a natural result of the increased scale height.

5. In these particular examples, the values of β at the maximum of electron production are comparable.

It should be stressed that these examples should not be taken as 'typical,' for the variations between individual records can be large.

The method of analysis seems capable of giving values for q_0 , β_0 , and H_0 within, say, 20 per cent. Errors due to the assumption of constant Γ are most important when the F_1 critical point goes to great virtual heights, i.e., at sunspot minimum summer or during storms. Errors due to the valley between E and F_1 , or alternatively to a separate source of F -region ionization, as well as errors due to overlap of the o and x traces on the ionogram, are most important when the virtual height record is very 'flat,' i.e., at sunspot maximum winter.

It is essential to use ionograms of the highest quality, avoiding those with sporadic E blanketing a part of the F region or with irregularities due to traveling disturbances. The effect of using average $n(h)$ data for deducing average values of the parameters has been considered; in the author's opinion, this is not advisable. The most accurate methods of $n(h)$ reduction at

present available are not fully adequate a problem, particularly in regard to getting accurate F -region profile when there is a valley between the E and F_1 layers.

The art is not yet in a state where the assumptions can be tested critically, but the over-all agreement between the observed $n(h)$ distributions and the theory based on the assumptions permits a measure of confidence in the values deduced for the ionospheric parameters.

Acknowledgments. This work is part of a research program of the Geophysical Observatory of the New Zealand Department of Scientific and Industrial Research (Superintendent, J. W. H. Ley). It has benefited greatly by discussion with C. H. Cummack. The assistance of members of the technical staff is much appreciated.

REFERENCES

- Hirsh, A. J., *J. Atmospheric and Terrestrial Physics*, **17**, 86, 1959.
- King, G. A. M., *J. Atmospheric and Terrestrial Physics*, **5**, 245, 1954.
- King, G. A. M., and F. E. Roach, *J. Nuclear Energy, B*, **1**, 129, 1961.
- Nicolet, M., *J. Atmospheric and Terrestrial Physics*, **1**, 141, 1951.
- Ratcliffe, J. A., *J. Atmospheric and Terrestrial Physics*, **8**, 260, 1956.
- Yonezawa, T., H. Takahashi, and Y. Arai, *Radio Research Labs.*, **6**, 21, 1959.

(Manuscript received March 28, 1961;
revised June 6, 1961.)

Effect of Electron-Ion Collisions in the F Region of the Ionosphere on the Absorption of Cosmic Radio Noise at 25 Mc/s at Ahmedabad

Changes in Absorption Associated with Magnetic Storms

K. R. RAMANATHAN, R. V. BHONSLE AND S. S. DEGAONKAR

Physical Research Laboratory, Ahmedabad-9 India

Abstract. Measurements of cosmic radio noise at 25 Mc/s, being made at the Physical Research Laboratory, Ahmedabad, since March 1957, have shown much larger values of absorption than those observed by Shain and Mitra in Australia. This fact, together with the empirically known dependence of the absorption on the critical frequency of the F region, and the effect of magnetic storms on the absorption found from the Ahmedabad observations have led us to examine the different possible parameters that may affect cosmic-noise absorption. It is found that electron-ion collisions in the F region both below and above the level of maximum electron density contribute in a substantial way to the absorption of cosmic radio noise. The values of hourly absorption due to collisions of electrons with neutral particles and with ions have been calculated for a period of 6 days in August–September 1957, when there were three magnetic storms. The results obtained show a depletion of electrons above F maximum on the day following the commencement of the magnetic storm and a refilling on later days. The results are discussed in relation to findings from satellite observations about particle fluxes in the Van Allen belts during magnetic storms.

Introduction. The cosmic-radio-noise method of measuring ionospheric absorption was used by Mitra and Shain [1953] in Australia at 18.3 Mc/s and by Bhonsle and Ramanathan [1958] at Ahmedabad at 25 Mc/s. Mitra and Shain found that the total ionospheric attenuation could be divided into two parts, one of which was attributed to absorption in the D region of the ionosphere and the other mainly to absorption in the F region. Applying Mitra and Shain's method of analysis to the Ahmedabad observations, it was noticed that, although the operating frequency of 25 Mc/s was higher than that used in Australia, the magnitude of the total attenuation observed at Ahmedabad during the period 1957 to 1959 was much higher than that observed in Australia (in 1950). This high value of attenuation at Ahmedabad was partly expected, since Ahmedabad is situated near the latitude of the F_2 ionization. Examination of the mean diurnal and seasonal variations of total attenuation at Ahmedabad showed the following features:

3. Whenever the second hump occurred, it did so in the early part of the night. It was most frequent in equinoxes and some winters.

On examining the Ahmedabad data, it was believed that it would be better to divide the total attenuation into a daytime component 'symmetrical' about noon, and a residual 'non-symmetrical' component without any implied commitment regarding the region of the ionosphere in which the absorption might originate. Unambiguous correlation of the nonsymmetrical component with the critical frequency f_oF_2 of the F_2 layer showed that its major part originated in the F region, although there might be residual effects originating well above the level of maximum electron density.

Bhonsle [1960] extended the analysis to cover the Ahmedabad data for the whole period of the IGY and IGC, and confirmed the trends of diurnal and seasonal variations of the total attenuation described above. The monthly mean values of the total attenuation exhibited maxima in equinoxes and minima in summer and winter. This feature was repeated from year to year. The seasonal variations of the monthly mean values of f_oF_2 were also similar, and analysis showed a close relationship between them. Figure 1 gives a mass plot of the mean

The monthly mean diurnal curves of the total attenuation exhibited either one or two humps, The first hump occurred in most of the months between 14 and 16 hours local time,

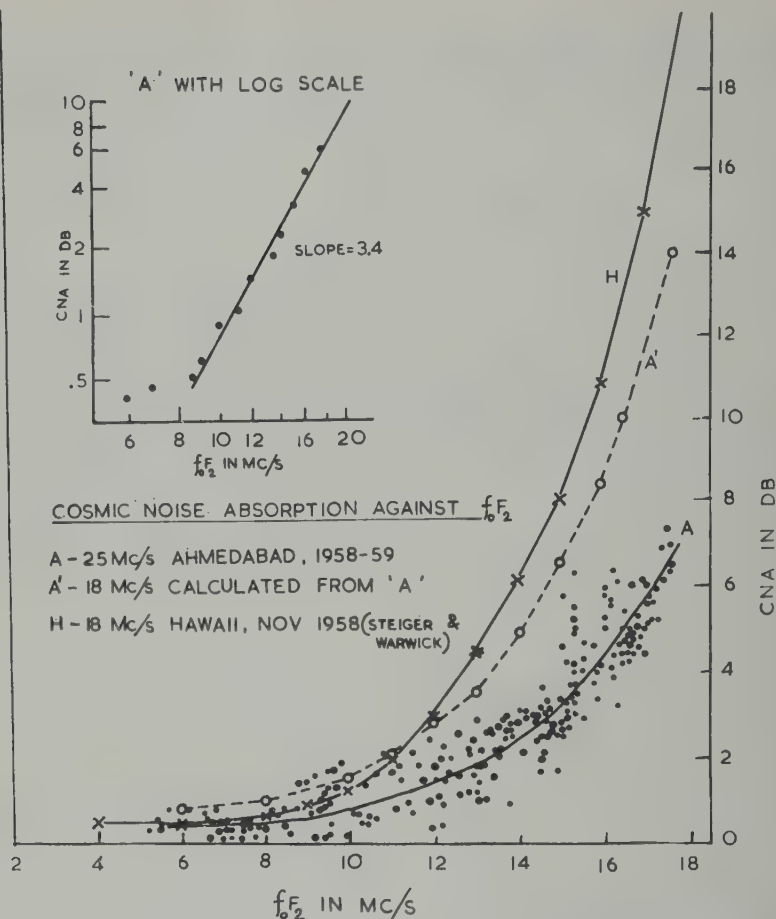


Fig. 1. Cosmic-noise absorption against $f_o F_2$ at Ahmedabad and Hawaii.

hourly values of cosmic-noise absorption at 25 Mc/s at Ahmedabad in each month of 1958 and 1959. The mean curve is also drawn with logarithmic scales on the left-hand top corner of Figure 1. It is interesting to note that there is an approximately linear relationship between $\log f_o F_2$ and the logarithm of attenuation, when $f_o F_2$ is greater than 8 Mc/s. Under this condition, the attenuation can be expressed in the form

$$\text{Attenuation}^1 = K + \lambda(f_o F_2 - 8)^{3.4}$$

¹ The significance of λ and of the exponent 3.4 is discussed in the Appendix. The mean curves of attenuation at 18 Mc/s as observed by Steiger and Warwick [1961] at Hawaii and as calculated from the Ahmedabad curve for 25 Mc/s assuming an inverse square relationship with $f_o F_2$ are also drawn in the same figure.

In a study of the effect of magnetic storms on the total attenuation of cosmic noise, Bhonsle and Ramanathan [1960] found that the attenuation became abnormally low on the first 2 days after the sudden commencement of a storm but increased to values above normal on the third day after the commencement. The observed changes in attenuation associated with magnetic storms were tentatively explained by being due to changes in $f_o F_2$ and in F_{min} .

Many years ago, Cowling [1945] showed that when an electron moves through an atmosphere containing uncharged atoms and molecules, the effective cross section for collision between an electron and an ion is much larger than that for a collision between an electron and a neutral particle. In calculating the collision frequencies of elec-

the ionosphere, we have to take into account only collisions between electrons and neutral particles but also collisions of electrons with positive ions. Whereas in the *D* and *E* regions of the ionosphere the former predominates, the situation is reversed in the *F* region. The electron collisions in the higher regions of the ionosphere become particularly important in the regions where the maximum electron densities are high.

In this paper, we first calculate the collision frequencies of electrons at different levels in the ionosphere up to the level of maximum electron density at each hour of a period of 6 months from August 31 to September 5, 1957. In making this calculation, the collisions of electrons with neutral atoms or molecules and with positive ions are both considered. The tables of *Onozawa* [1960] and of *Nicolet* [1959] have been used for the physical properties of the upper atmosphere. Electron densities at different heights over Ahmedabad have been calculated from the ionospheric vertical-sounding records. At levels above $h_{\max} F_2$, two extreme models of electron distribution, one as given by *Alpert*, *Chudensenko*, and *Shapiro* [1958], and the other by *Kazantsev* [1959], have been used. (See also *Garriott* [1960]).

Calculation of absorption in different regions. Following *Chapman and Little* [1957], the absorptive absorption of radio waves of high frequency f Mc/s in the vertical direction can be written in the form

$$D = A/f^2 \dots, \dots, \dots \quad (1)$$

where D is the absorption in decibels and

$$A = 1.17 \times 10^{-14} \times \int_0^H N_e \nu \, dh \dots, \dots, \dots \quad (2)$$

where N_e = the number of electrons per cc.

ν = the collision frequency of electrons (number of collisions per second).

dh = an element of path length along the vertical.

H = the maximum height to which integration is taken.

In equation 1, the longitudinal component of magnetic frequency f_L is neglected in com-

parison with f , and ν in comparison with $6\pi \times 10^6$.

For the collision frequency of electrons with neutral atoms and molecules in the atmosphere, *Nicolet* [1959] has given the formula

$$\nu_{em} = 5.4 \times 10^{-10} N_m T^{1/2} \dots, \dots, \dots \quad (3)$$

where ν_{em} is the frequency of collision of electrons with neutral particles, and N_m is the total number density of neutral particles (N_2, O_2, O , etc.), and T is the temperature of the atmosphere in degrees Kelvin.

Cowling's expression for the frequency of collision of electrons with positive ions as revised by *Nicolet* [1959] is

$$\nu_{ei} = (34 + 4.18$$

$$\cdot \log (T^3/N_e)) N_i / T^{3/2} \dots, \dots, \dots \quad (4)$$

If we assume that in the ionosphere most of the positively charged particles are singly ionized,

$$N_i \approx N_e$$

TABLE 1

Sample table showing the variation N_m , N_e , ν_{em} , ν_{ei} , and $N_e \nu$ with true height at 15 hours on September 7, 1957, at Ahmedabad. Following *Chapman*, the notation 2.2^6 is used to denote 2.2×10^6 .

Height, km	T , °K	N_m	N_e	ν_{em}	ν_{ei}	$N_e \nu$
70	210		4.2 ²	1.1 ⁷		4.2 ⁹
80	197		4.2 ²	2.2 ⁶		8.3 ⁹
90	197		1.1 ⁴	4.9 ⁶	1.9 ²	5.0 ⁹
100	210	1.5 ¹³	3.0 ⁴	1.1 ⁵	4.4 ²	3.0 ⁹
120	380	9.6 ¹¹	1.9 ⁵	1.0 ⁴	1.1 ³	2.1 ⁹
140	513	2.2 ¹¹	2.2 ⁵	2.7 ³	8.8 ²	7.9 ⁸
160	616	7.9 ¹⁰	2.7 ⁵	1.1 ³	8.1 ²	5.2 ⁸
180	702	3.5 ¹⁰	3.6 ⁵	5.1 ²	8.4 ²	4.7 ⁸
200	775	1.8 ¹⁰	4.5 ⁵	2.7 ²	9.1 ²	5.3 ⁸
250	909	4.9 ⁹	7.9 ⁵	8.0 ¹	1.3 ³	1.3 ⁹
300	1038	1.1 ⁹	1.4 ⁶	2.7 ¹	2.0 ³	2.8 ⁹
350	1140	6.5 ⁸	2.9 ⁶	1.2 ¹	3.7 ³	1.1 ¹⁰
400	1220	2.9 ⁸	3.3 ⁶	5.4 ⁰	3.4 ³	1.1 ¹⁰
450	1400	1.1 ⁸	3.1 ⁶		2.5 ³	7.8 ⁹
500	1580	6.1 ⁷	2.6 ⁶		1.6 ³	4.2 ⁹
600	1580	1.9 ⁷	1.8 ⁶		1.1 ³	2.0 ⁹
700	1580	6.9 ⁸	1.3 ⁶		8.0 ²	1.0 ⁹
800	1580		9.3 ⁵		6.0 ²	5.6 ⁸
900	1580		7.0 ⁵		4.6 ²	3.2 ⁸

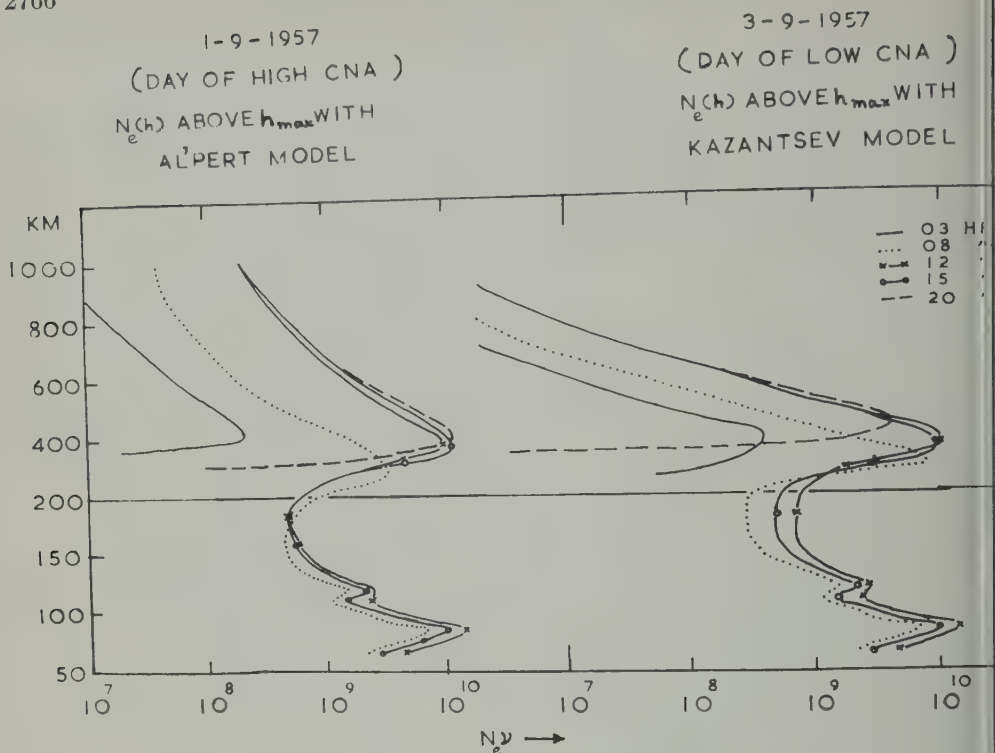


Fig. 2. $N_e \nu$ against height over Ahmedabad at different hours on September 1 and 3, 1957.

For the same electron density, the collision frequency decreases with increase of temperature.

To calculate A in equation 2, ν should be taken to be the sum of ν_{em} and ν_{ei} . The collision cross section of an electron with a positive ion is much larger than that with a neutral atom or molecule. So, while in the D and E regions of the ionosphere, ν_{em} is more important than ν_{ei} , in the F region and above, with decreasing neutral particle density and increasing electron density, ν_{ei} becomes more important. During daytime, ν_{ei} becomes equal to ν_{em} at about 170 km.

The values of N_e for zenith sun at levels below 90 km were taken from Nicolet and Aikin [1960] corresponding to disturbed solar conditions. Between 90 and 120 km, they were interpolated from the values at 90 and 120 km. Above 120 km to the level of maximum electron density in F_2 the electron densities were calculated for true heights from the ionospheric sounding records using Schmerling's coefficients. For still higher levels, the values of electron density were estimated on 2 days according to two extreme models, one given by Alpert and

others and the other by Kazantsev. Both were based on observations of radio waves received from satellites. From these a number densities of particles at different heights as given by Nicolet (up to 120 km), Yonezawa (120 to 420 km), the values of ν_{ei} and $N_e \nu$ were calculated. For levels above 500 km, a uniform temperature of 1580°K was assumed for calculating ν_{ei} . The manner in which ν_{em} and ν_{ei} vary with height is shown in sample Table 1, and the variation of $N_e \nu$ with height in Figure 2. The different curves correspond to 3, 8, 12, 15, and 20 hours on September 1 and 3, 1957, over Ahmedabad. The absorption due to electron-neutral particle collisions and electron-ion collisions were separately calculated. The product $N_e \nu_{em}$ attains a maximum at about 85 km and falls off with increasing height. Above 240 km, only electron-ion collisions are important. The absorption is maximum at the level of maximum ionization of the F region.

Numerical integration of $N_e \nu_{dh}$ was carried out by dividing the ionosphere into 5-km

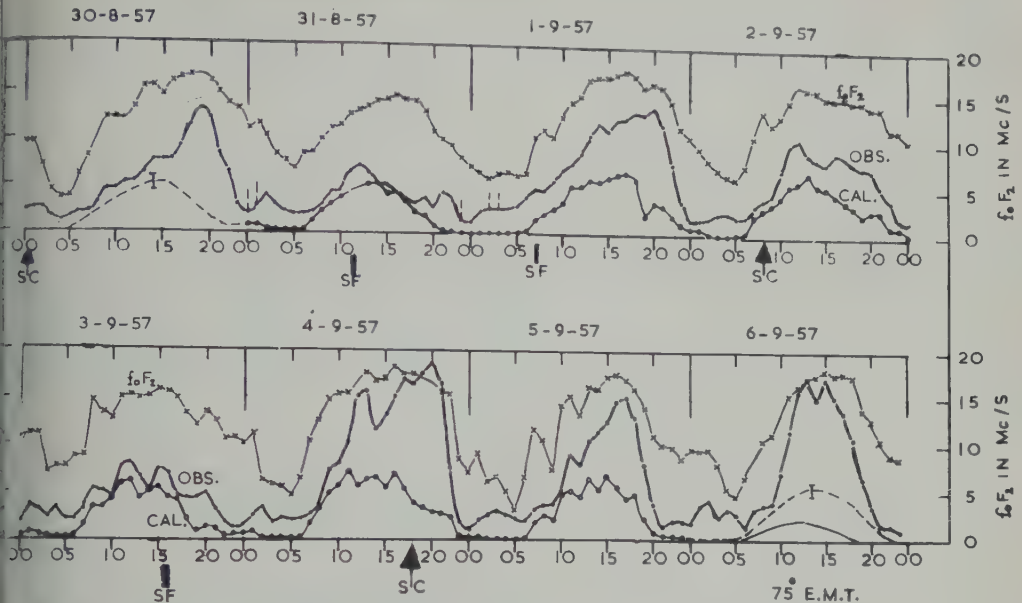


Fig. 3. Observed and calculated values of cosmic-noise absorption at 25 Mc/s and f_oF_2 at Ahmedabad for the period August 30 to September 6, 1957. OBS, observed cosmic-radio-noise absorption at 25 Mc/s at Ahmedabad; CAL, calculated cosmic-radio-noise absorption at 25 Mc/s from $N_e(h)$ profiles at Ahmedabad; f_oF_2 , F_2 critical frequency for ordinary waves.

between 65 and 120 km, 20-km intervals between 120 km to h_mF_2 , and 50-km intervals between h_mF_2 and 1000 km. The calculated absorption in the region 65 to 120 km was about 1 db for zenith sun. Absorptions in this region at other values of solar zenith angles were calculated on the assumption that it obeys a \sec^2 law, where $n = 0.75$. Also for this region, the same average electron-density distributions depending on local time were assumed for all days under investigation.

Comparison of observed values of total absorption and calculated values of absorption below h_mF_2 . The total ionospheric attenuations measured with cosmic radio noise on 25 Mc/s at Ahmedabad for the period August 30 to September 6, 1957, are plotted in Figure 3. The calculated values of ionospheric absorption below h_mF_2 and the values of f_oF_2 are also plotted in the same figure for comparison. There occurred SC-type magnetic storms during this period. Their times of commencement are marked in the figure. The measured total attenuations exhibit a diurnal variation with large day-to-day changes. The calculated absorptions in the region below h_mF_2 show a maxima of 2

to 3 db between 13 to 16 hours local time on all these days, and smaller values of absorption in the morning and evening hours. The measured total attenuations are always higher than the corresponding calculated values of absorption for the region below h_mF_2 . No doubt this is largely due to the contribution to absorption from regions above F_2 peak.

In Figure 4 are plotted for September 1, 1957, in addition to the measured and calculated absorptions below h_mF_2 , the calculated hourly absorptions up to 1000 km, $A(A)$, and $A(K)$, assuming Al'pert's and Kazantsev's models of electron-density distribution above $N_{max}F$. On this day, absorptions calculated on Al'pert's model $A(A)$ were generally in good agreement with observation, though there is some disagreement between the observed and calculated curves in the evening hours. Kazantsev's model gave too small values. On September 3, 1957, a day on which the measured attenuation was abnormally low, the absorptions calculated on Kazantsev's model gave a reasonable fit, but there was some discrepancy in the evening hours. The results can be interpreted to mean that the total number of electrons over

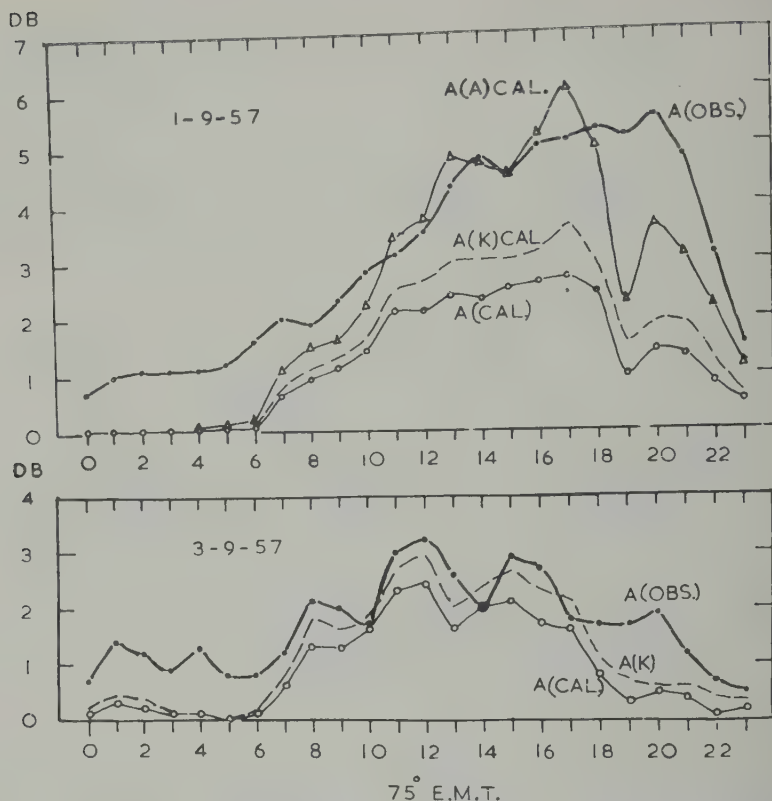


Fig. 4. Observed and calculated absorptions on September 1 and 3, 1957. A(OBS.), observed cosmic-noise-absorption on 25 Mc/s at Ahmedabad; A(CAL.), calculated cosmic-noise-absorption in the region 65 km to $h_m F_2$; A(K)CAL., calculated cosmic-noise-absorption from 65–1000 km using Kazantsev's model of electron-density distribution above $h_m F$; A(A)CAL., calculated cosmic-noise-absorption from 65–1000 km using Al'pert's model of electron-density distribution above $h_m F$.

unit area above $N_{\max} F$ was much less on August 3, 1957, than on September 1, 1957.

The hourly differences between the observed total attenuations and the corresponding calculated values of absorption below $h_m F$ on 8 days from August 30 through September 6, 1957, are plotted in Figure 5. This difference gives us an estimate of absorption occurring in the region above $h_m F$. It appears that (1) large day-to-day changes in absorption occur above $h_m F$ in magnetically disturbed periods; (2) the absorption above $h_m F$ is very much reduced on the second day after the commencement of a magnetic storm; and (3) on the following day the absorption rises up above its prestorm value.

The evidence is strong that the total number of electrons in the atmosphere above $h_m F_2$ is

largely depleted in the closing phase of a magnetic storm, and that it is more than restored at the end of the storm.

Discussion. In the above calculations of absorption, only the nondeviative part of the absorption due to electron collisions has been taken into account. Deviative absorption and diffraction of aerial aperture can also cause absorption. Both these depend on the nearest critical frequency (25 Mc/s) to the critical frequency of the F_2 layer. Since the nearest critical frequency over Ahmedabad rarely exceeds 18 to 19 Mc/s, and since the aerial aperture of the aerial was limited to 40° in the vertical, it can be shown that both of these two can be neglected. It has been pointed out [Ramanathan and Bhonsle

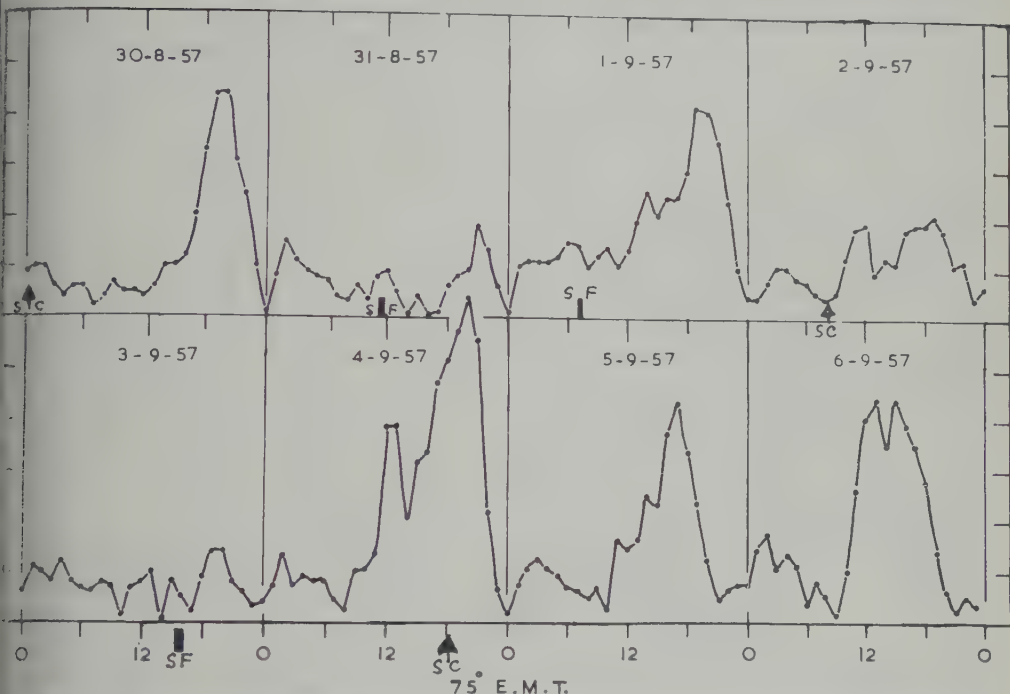


Fig. 5. Difference between observed cosmic-radio-noise absorption on 25 Mc/s and the absorption calculated for the region below $h_m F$ from $N_e(h)$ profiles at Ahmedabad for the period August 30 to September 6, 1957.

F scatter may cause an increase in the variation. Examination of day-to-day records shows, however, that F scatter can cause either enhancement or a reduction in the intensity of cosmic noise, depending on the sidereal time and the phase of the daily variation of intensity. The question is under further examination and results will be reported separately.

Figures 3 and 4 show an additional diurnal variation in absorption in the late evening hours after sunset. This may be due to the following reasons: (1) the distribution of electron density at height above F_2 peak may vary with the time of the day. It is probable that at the time of sunset F_2 has begun to decrease in the evening. After reaching its peak, there is a delay in the decrease of electron density at levels above the F_2 peak, and (2) there may be a marked diurnal variation of temperature above 300 km in the region where electron-ion collisions contribute toward the attenuation of cosmic radio waves. In fact, there is evidence from rocket and satellite data for both these. Nisbet [1960], in a

paper on electron-density distribution in the upper ionosphere, has collected together the available rocket and satellite data. They indicate that above $f_o F_2$ peak high electron densities continue to persist in the late evening hours and that they decrease with height much more rapidly in the second half of the night and in the morning. Ross [1960] has shown from the Doppler frequency measurements of satellite 19588₂ that on the average, there is a steady increase in total electron content between the F_2 peak and the satellite (height approximately 950 km) from sunrise to sunset, even though the electron content below $h_{max} F$ had begun to show a decrease by sunset. From the data of deceleration of satellites as analyzed by Jacchia, Martin, and Priester, and others, Nicolet [1960] concludes that between 300 and 800 km there is probably a cooling during the night of the order of 500°K. Since ν_e is proportional to $T^{-3/2}$, a cooling of the upper part of the ionosphere of this order will cause an increase in the collision frequency of electrons in the ratio of 1.5 to 1.6 if the elec-

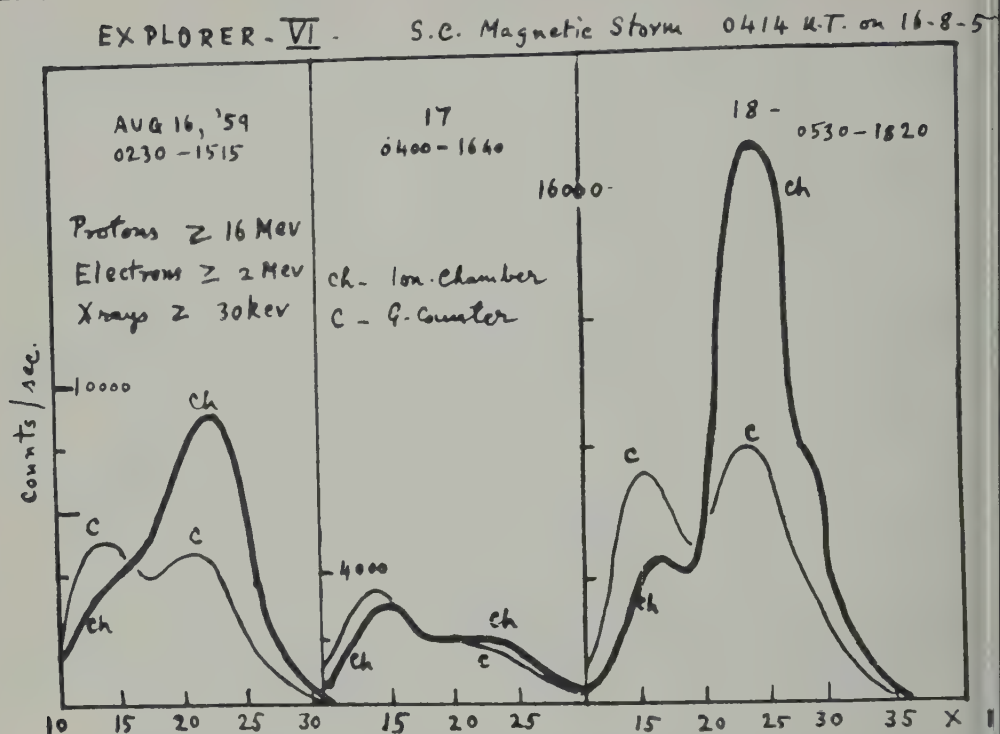


Fig. 6. Particle fluxes in Van Allen radiation belts on August 16, 17, and 18, 1959 (observations with Geiger counter and ion chamber in Explorer VI, after Arnoldy, Hoffman, and Winckler).

tron density remains the same. Thus it appears that the post-sunset increase in radio noise absorption is due to the combined effect of a decrease in the temperature of the outer ionosphere, together with a lag in the rate of attachment of electrons with increasing height.

The effect of F scatter on the attenuation is quite considerable on some nights and will be dealt with separately.

It is interesting to compare these changes in cosmic noise absorption during and after magnetic storms with the changes in particle fluxes that have recently been observed in Van Allen belts during such storms. Figure 6 shows the observations made by Arnoldy, Hoffman, and Winckler [1960] during a magnetic storm that commenced at 0414 UT on August 16, 1959, with instruments carried in Explorer VI. The observations cover a period of 3 days, and the figure shows the counts per second recorded by an ion chamber and a Geiger counter carried in the satellite. The counts of the ion chamber are shown by thick lines and those of the Geiger

counter by thin lines. The remarkable feature brought out by the diagram is that on the second day of the storm, there was a large decrease in radiation in the outer Van Allen belt, and it recovered to more than its prestorm value by the third day (August 18). The changes in harder radiation recorded by the Geiger counter are comparatively smaller, though they are in the same direction. Considering the similarity between the changes in cosmic-radio-noise absorption and the flux of energetic particles in the Van Allen belts, it appears that during magnetic storms there is a large decrease in the number of electrons even in the outer ionosphere in low and middle latitudes during the early phase of a magnetic storm, and that at the end of the storm the outer region gets refilled with electrons. There are now quite a few examples of satellite observations that bear out this general conclusion. Whereas the particle counters carried in satellites measure only energetic particles with energies above certain limits, cosmic-radio-noise absorption indicates changes

total number of electrons including slow ions.

Acknowledgment. We wish to thank the Council of Scientific and Industrial Research, India, for financial assistance.

APPENDIX

According to equation 2,

$$A = 1.17 \times 10^{-14} \int_0^H N_e \nu \, dh$$

In the F layer, ν_{em} can be neglected in comparison with ν_{ei} , and

$$\nu_{ei} = (34 + 4.18 \log T^3/N_e) N_e / T^{3/2}$$

$$N_e = 1.24 \times 10^4 f^2$$

where f is the plasma frequency in megacycles per second, and

$$A = (34 + 4.18 \log T^3/N_e) N_e^2 / T^{3/2}$$

$$= 1.54 \times 10^8 (34 + 4.18 \log T^3/N_e) f^4 / T^{3/2}$$

In the region between 300 and 1000 km, the values under brackets can vary from 44 to 50. At the level of maximum electron density $f_o F_2$, and since for some distance both above and below that level the plasma frequency may be expected to vary with $f_o F_2$, and since according to our present understanding of temperatures in the F_2 peak, the temperature increases with height or remains steady at a higher value than at $h_m F_2$,

$$\int_{300 \text{ km}}^{1000 \text{ km}} N_e \nu_{ei} \, dh$$

may be expected to vary as some power of $f_o F_2$ where the exponent is a little smaller than 4.

REFERENCES

- Al'pert, Dobriakova, Chudensenko, and Shapiro, *Doklady Akad. Nauk SSR*, 120 (4), 743-746, 1958.
- Arnoldy, R., R. Hoffman, and J. R. Winckler, *Space Research*, First International Space Science Symposium, edited by H. Kallmann-Bijl, pp. 877-896, 1960.
- Bhonsle, R. V., Ph.D. thesis, Gujarat University, 1960.
- Bhonsle, R. V., and K. R. Ramanathan, *J. Sci. Industrial Research (India)*, 17A, 40, 1958.
- Bhonsle, R. V., and Ramanathan, K. R., *Planetary Space Sci.*, 2, 99, 1960.
- Chapman, S., and C. G. Little, *J. Atmospheric and Terrest. Phys.*, 10, 20, 1957.
- Cowling, T. G., *Proc. Roy Soc.*, 183A, 453, 1945.
- Garriott, O. K., *Space Research*, First International Space Science Symposium, edited by H. Kallmann Bijl, pp. 371-386, 1960.
- Kazantsev, A. N., *J. Planetary Space Sci.*, 1, 130-135, 1959.
- Mitra, A. P., and C. A. Shain, *J. Atmospheric and Terrest. Phys.*, 4, 204, 1953.
- Nicolet, M., *Physics of Fluids*, 2, 95, 1959.
- Nicolet, M., *Space Research*, First International Space Science Symposium, edited by H. Kallmann-Bijl, 46-89, 1960.
- Nicolet, M., and A. C. Aikin, *J. Geophys. Research*, 65, 1469-1483, 1960.
- Nisbet, J. S., *J. Geophys. Research*, 65, 2597-2599, 1960.
- Ramanathan, K. R., and R. V. Bhonsle, *J. Geophys. Research*, 1635-1637, 1959.
- Ross, W. J., *J. Geophysical Research*, 2607-2615, 1960.
- Steiger, W. R., and J. W. Warwick, *J. Geophys. Research*, 66, 57-66, 1961.
- Yonezawa, T., *J. Radio Research Lab. (Tokyo)*, 7, 69, 1960.

(Manuscript received May 26, 1961.)



Suggestion of a Second Isopycnic Level at 80 to 90 Kilometers over Churchill, Canada

ALLEN E. COLE

*Geophysics Research Directorate, Air Force Cambridge Research Laboratories
Bedford, Massachusetts*

Abstract. All available observations from rocket-borne instrumentation (falling spheres, grenades, pressure gages, etc.) have been used to examine the distribution of atmospheric density over Churchill, Canada. An observed reduction in seasonal and interdiurnal variability between 80 and 90 km strongly suggests the presence of a second isopycnic level in this region. Daily and seasonal density profiles appear to converge and cross at this altitude.

Introduction. Until recently, most meteorologists considered atmospheric density to be a derived quantity, secondary in importance to temperature and pressure for describing the earth's atmosphere. In the space age, however, atmospheric density has become increasingly important. Designers and engineers need information on the distribution of density to determine altitudes to predict the performance of missiles, satellites, and aircraft. Density is the defining parameter for model atmospheres above 8 km, because it is measured more accurately at high altitudes than either pressure or temperature.

The first systematic investigation of the combined effect of pressure and temperature on atmospheric density at various elevations above the earth's surface [Sen, 1924] was limited by available data to the first 10 km of the earth's atmosphere. Examining the seasonal changes in isopycnic surfaces (surfaces of constant density) at various slopes of various equisubstantial surfaces (isopycnic surfaces) from pole to equator, Sen found that as density increased below 8 km it decreased above 8 km, and, conversely, a decrease below 8 km accompanied an increase above 8 km. He concluded that below 8 km air density was controlled primarily by air temperature but above 8 km primarily by pressure changes.

Sen proposed the names *thermosphere* and *thermosphere* for the atmospheric layers below and above 8 km, respectively, and *thermopause* for the neutral layer at approximately 8 km, where he found that density remained constant during seasons and at all locations over the whole earth. These terms were never adopted, but

the isopycnic level at 8 km and the seasonal changes in the structure of the atmosphere which he described have been recognized by other researchers.

Humphreys [1929], DoPorto [1943], Morgan [1948], and others have extended Sen's study to altitudes up to 25 km. In response to an increasing requirement from designers of re-entry vehicles, Sissenwine, Ripley, and Cole [1958] investigated seasonal and interdiurnal density variation in subtropical, temperate, subarctic, and arctic regions. They presented seasonal means and standard deviations of densities at levels 2 km apart up to approximately 28 km, and matrices of correlation coefficients of densities at pairs of levels. The isopycnic level was found around 8 km at all locations, and a second such level near 26 km was suggested.

More recent data show a reduction in interdiurnal and seasonal variability between 24 and 26 km, but not to the extent that an isopycnic surface can be said to occur at 26 km. This pseudo-isopycnic level is prominent on all latitudinal and longitudinal cross sections, lying between regions where seasonal and interdiurnal density variability is relatively large.

Information on the variability of density at levels above 30 km is derived primarily from rocket observations at Churchill, Canada (59°N), or at missile ranges between 30° and 38°N in the United States. These data indicate that the seasonal and interdiurnal variations of atmospheric density between 30 and 90 km are greatest at high latitudes.

Sufficient observations are not available at these altitudes for the preparation of statistical

matrices of seasonal means, standard deviations, and correlation coefficients. However, available data for Churchill provide an indication of the magnitude of seasonal and interdiurnal density variations for altitudes between 30 and 90 km.

Data used to examine the seasonal and interdiurnal density variability at Churchill included radiosonde observations from 1956 through 1959 up to heights of 30 km and all available observations from rocket-borne instrumentation (falling spheres, grenades, pressure gages, etc.) between 30 and 90 km. Individual density soundings are available from 17 winter (Nov.-March), 12 summer (June-Aug.), and 3 autumn (Oct.) rocket flights north of latitude 50. Approximately two dozen temperature observations to 40 and 45 km were available from meteorological rockets fired at Churchill during 1960.

Procedure. Mean temperature-height profiles (Fig. 1) for Churchill from the surface to 90 km were prepared for winter (Nov.-March) and summer (June-Sept.) seasons. Mean density profiles were computed for each season from these temperature-height profiles, using the perfect gas law and the hydrostatic equation. These are plotted in Figure 2 as percentage departures from a proposed (March 1961) Revised U. S. Standard Atmosphere and Extension [Cole, Court, and Kantor, 1961], intended to represent mean annual conditions at 45°N. Individual density observations are also shown as percentage departures from this proposed standard.

The individual density observations between 30 and 50 km vary from approximately 45 per cent less than the proposed standard in winter to 35 per cent greater than the standard in summer. Between 50 and 80 km the observed range is slightly greater: -57 to +35 per cent. At 85 km the observed annual range decreases; most of the individual soundings for which density was less than the standard between 30 and 80 km approach or become greater than the standard at 90 km, and vice versa.

Consequently, the mean season profiles cross in this vicinity; the mean winter profile becomes greater than standard while the summer profile approaches the standard. This phenomenon was also found at approximately 80 to 90 km in the preparation of model atmospheres; increases in temperature at 50 km result in lower densities at this level and higher densities above 80 km.

Reduction in temperature at 50 km results in an opposite effect.

Possible isopycnic level. The observed variation in seasonal and interdiurnal variability suggests the presence of a second isopycnic level between 80 and 90 km over Churchill. Unfortunately, relatively few density observations are available for this altitude, and the accuracy of these measurements, most of which were north of 50°, is often questionable. This is based primarily on measurements from different sources: falling spheres, rocket pressure gages, and grenades. Individual ascent profiles have been summarized by Quiroz [1961].

For falling-sphere densities, Jones and Jones [1961] indicate an accuracy of ± 5 per cent between 80 and 90 km. Densities derived from grenade data are estimated to have errors of less than 10 per cent [Nordberg and Stroud, 1961]. N. W. Jones and D. R. Taesch (private communication) indicate that their smoothed densities, derived from pressure measurements, are accurate within ± 10 per cent between 80 and 90 km.

Even allowing for ± 10 per cent error in individual density observations plotted in Figure 2 tend to converge above 80 km. These densities approach or become greater than the proposed standard in practically all cases. In one of the two summer observations the density at altitude is less than standard.

Sufficient density observations are not available to determine the latitudinal extent of the surface of minimum density variability. A study by Nordberg and Stroud [1961], however, tends to support the presence of such a level at other latitudes. Their mean winter and summer density profiles for Churchill, using only falling-sphere data, converge above 80 km and appear to level off near 90 km. Their mean annual density profile for low latitudes, based on grenade data from White Sands, Guam, and Woomera, Australia, departs from both the Churchill winter and summer density profiles below 80 km. Above 80 km it converges with them, indicating a latitudinal and seasonal density change between 80 and 90 km.

In a related study at higher altitudes, Jones [1960] postulates a level of pressure and temperature equalization at 200 km over the Arctic. The presence of such a level is supported by theoretical considerations rather than by observed data. LaGow, Horowitz, and Allen

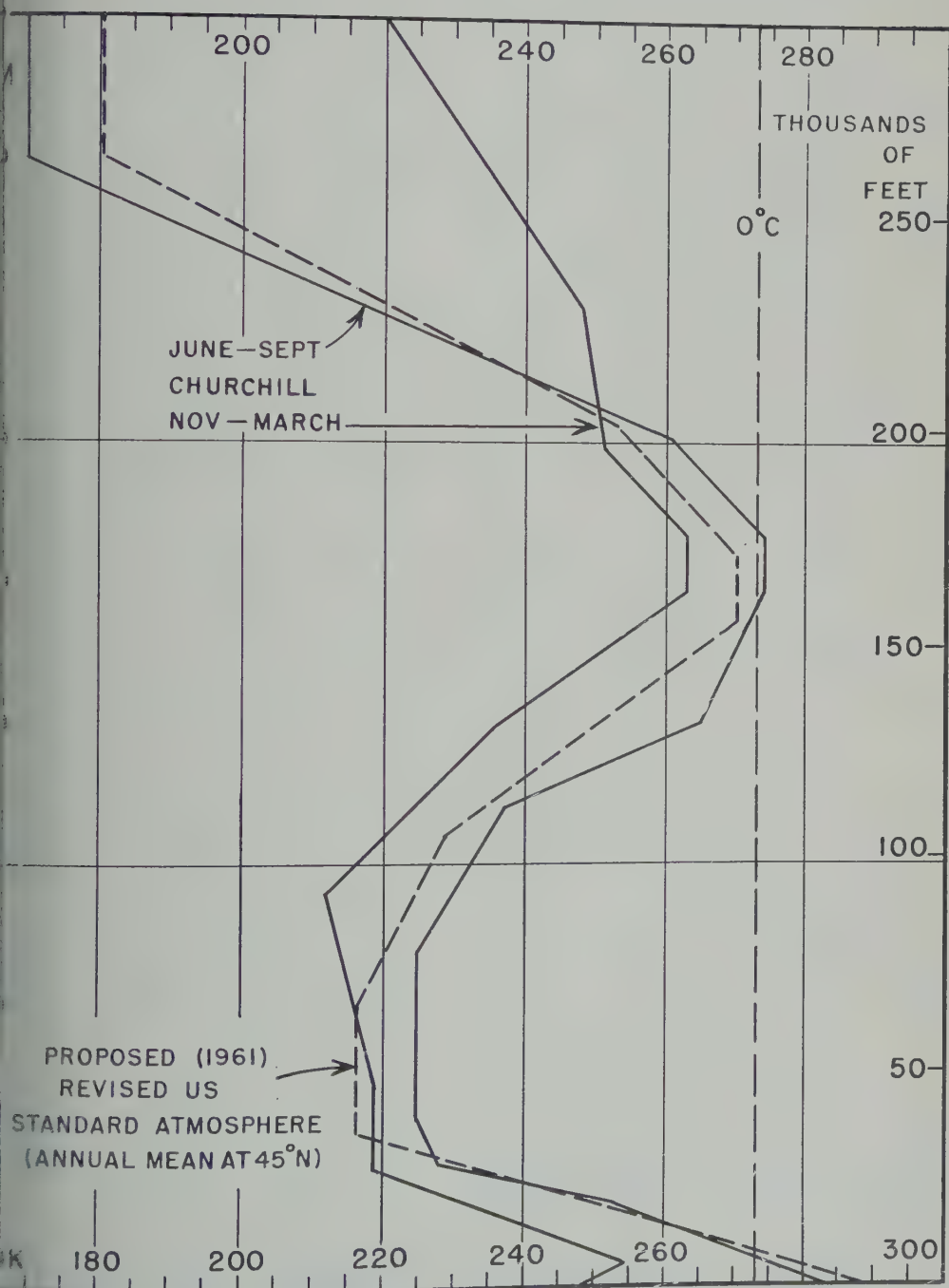


Fig. 1. Mean seasonal temperatures over Churchill and proposed mid-latitude standard.

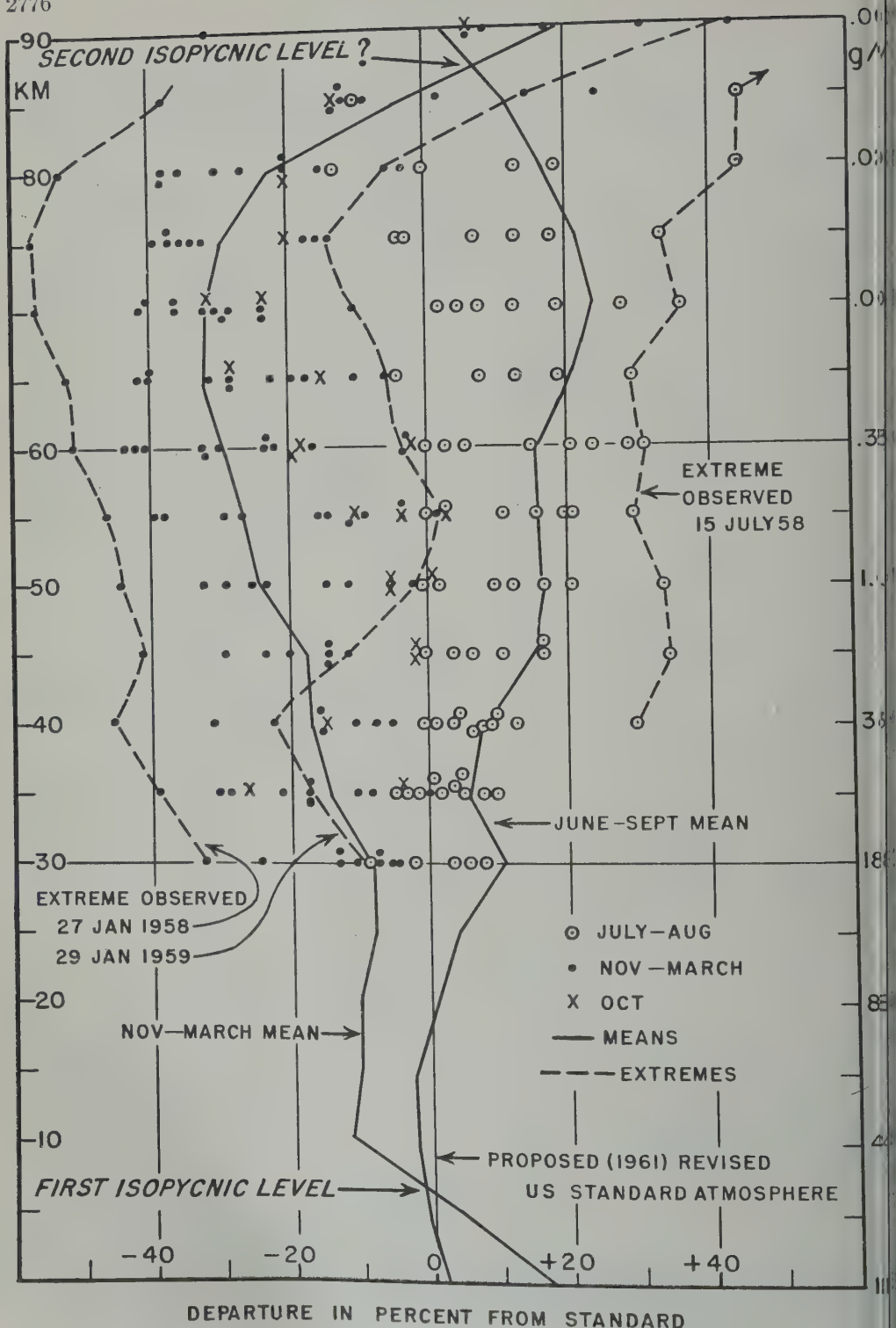


Fig. 2. Percentage departure of Churchill mean seasonal density profiles and observed density from proposed mid-latitude standard. The two extremes were observed in January 1958; 1959 a misprint.

indicate seasonal and latitudinal variations in the factor of 5 in observed densities in this region.

Inter. The two extreme individual density profiles observed at Churchill are shown in Figure 2. The densities of the extreme winter profile, observed on January 27, 1958, by the falling-sphere technique, are 30 to 57 per cent greater than the proposed standard at all levels between 30 and 90 km. Two days later, on January 29, an observation, also shown in Figure 1, indicated that the density at 50 km increased by roughly 79 per cent.

At this level, therefore, density increased from approximately 20 per cent less to 25 per cent greater than the winter mean for Churchill in the winter period. Similar changes in the density profile were caused by a sudden warming in the stratosphere, described in detail by Jones, Peter-Schafer, and Schultz [1959] in a discussion of falling-sphere observations at Churchill.

How often sudden warmings occur at 50 km is not known. Scherhag [1952], Craig and Hering [1954], and McClain [1961] describe wintertime warmings which have been observed by balloon-borne instruments in recent years; all of them may or may not have been associated with stratospheric warmings at higher altitudes.

Summer. The extreme summer profile, based on pressure measurements, indicates densities approximately 35 per cent greater than the proposed standard between 40 and 75 km. Above 75 km, the differences between the proposed standard and the observed densities become smaller, contrary to the pattern observed in most of the other density profiles.

Consequently, the accuracy of the upper portion of this profile seems doubtful; higher densities in the region between 40 and 80 km would result in a more rapid decrease in pressure with altitude. An atmosphere in hydrostatic equilibrium must satisfy the differential equation $\rho^{-1} d\rho/dz = -g/p$, where ρ = density, p = pressure, g = acceleration of gravity, and z = geometric altitude. This requires a decrease in density at higher level if the atmosphere obeys the perfect gas law, $\rho = pM/R^*T$, where R^* = universal gas constant, T = temperature, and M = molecular weight of air. The level of sharp density decrease for such soundings appears to be in the region between 80 and 85 km.

Conclusions. The following conclusions can be drawn from the data presented:

1. Available density observations for Churchill suggest a surface of minimum density variability between 80 and 90 km. Additional observations, however, are required over Churchill and at other latitudes to determine whether this is a true isopycnic level or merely a region of reduced variability similar to that near 26 km.

2. The correlation between densities at levels between 30 and 80 km in individual profiles seems relatively high. Large departures from standard at one level are normally accompanied by similar departures of the same sign at other levels.

3. The observed day-to-day variability increases with altitude, from -45 to $+35$ per cent at 40 km to -57 to $+35$ per cent at 75 km. The observed day-to-day variability around mean seasonal (winter and summer) density profiles for Churchill is roughly ± 25 per cent.

Acknowledgments. The data used were obtained from many sources and represent the efforts of many persons. Recomputed and smoothed density data provided in unpublished form by N. W. Spencer and D. R. Taesch, NASA, and L. M. Jones, University of Michigan, are gratefully acknowledged. Arnold Court provided helpful discussions and suggestions.

REFERENCES

- Cole, A. E., A. Court, and A. J. Kantor, Standard atmosphere revision to 90 km, *Interim Notes on Atmospheric Properties*, no. 7 (revision 2), CRZWD, Air Force Cambridge Research Laboratories, Bedford, Mass., 1961.
- Craig, R. A., and W. S. Hering, The stratospheric warming of January-February 1957, *J. Meteorol.*, 16, 91-107, 1959.
- DoPorto, M., The computation of atmospheric pressure at 8 km level of constant air density, *Tech. Note 1*, Meteorological Service, Dublin, Ireland, 1943.
- Humphreys, W. J., *Physics of the Air*, McGraw-Hill Book Co., 2nd ed., 76-77, 1929.
- Johnson, F. S., Pressure and temperature equalization at 200-km altitude, *J. Geophys. Research*, 65, 2227-2232, 1960.
- Jones, L. M., and J. W. Peterson, Upper air densities and temperatures measured by the falling-sphere method, *Rept. 03553-5-T*, Office of University of Michigan Research Administration, 1961.
- Jones, L. M., J. W. Peterson, E. J. Schafer, and H. F. Schultz, Upper-air density and temperature: Some variations and an abrupt warming in the mesosphere, *J. Geophys. Research*, 64, 2331-2340, 1959.

- LaGow, H. E., R. Horowitz, J. Ainsworth, Results of IGY atmospheric density measurements above Fort Churchill, *Space Research, Proc. First Intern. Space Sci. Symposium*, Nice, 164-174, 1960.
- McClain, E. P., Thermal conditions in the arctic stratosphere near 80 W in January, *J. Meteorol.*, 17, 383-389, 1961.
- Morgan, W. A., A statistical analysis of the comparative accuracy of estimates of wind at 1 km using surface charts and 8 km using isopycnic charts, *Tech. Note 8*, Meteorological Service, Dublin, Ireland, 1948.
- Nordberg, W., and W. G. Stroud, Seasonal, latitudinal and diurnal variations in the upper atmosphere, *NASA Tech. Note D-703*, 1961.
- Quiroz, R. S., Air density profiles for the atmosphere between 30 and 80 kilometers, *Tech. Note 150*, Air Weather Service, USAF, 1961.
- Scherhag, R., Die Explosionartigen Stratosphären-erwärmungen des Spätwinter 1951/1952, *deut. Wetterdienstes in der U.S. Zone*, 38, 1952.
- Sen, S. N., On the distribution of air density the globe, *Quart. J. Roy. Meteorol. Soc.*, 50, 1924.
- Sissenwine, N., W. S. Ripley, and A. E. Cole, Behavior of atmospheric density profiles, *Air Force Surveys in Geophysics*, no. 109, Air Force Cambridge Research Laboratories, Bedford, 1958.
- Stroud, W. G., and W. Nordberg, W. R. Barber, F. L. Bartman, and P. Titus, Rocket-g measurements of temperature and winds in the mesosphere over Churchill, Canada, *J. Geophys. Research*, 65, 2307-2323, 1960.

(Manuscript received April 24, 1961;
revised June 6, 1961.)

A Model of Solar-Flare-Induced Ionization in the *D* Region

R. C. WHITTEN AND I. G. POPPOFF

Stanford Research Institute, Menlo Park, California

Abstract. With data from the Naval Research Laboratory Sunflare II rocket observation of a class 2⁺ solar flare that occurred on August 31, 1959, and the 18-Mc cosmic noise intensity record obtained during the flare, a model of flare-induced ionization in the *D* region was constructed, and a value for the dissociative recombination coefficient obtained. An experiment to determine the validity of the model and to obtain more precise values of the recombination coefficients is suggested.

Introduction. It has long been known that sudden ionospheric disturbances (SID), as short-wave fade-out, sudden cosmic absorption, sudden enhancement of atmospheres, etc., are manifestations of enhanced ionization in the *D*-layer which in turn is associated with the occurrence of a solar flare. Recently, however, has the nature of the ionizing portion of the solar flare radiation come into question. Measurements [Friedman, Chubb, Kupperian, and Lindsay 1958; Chubb, Friedman, and Kreplin, 1960a; Chubb, and others, 1960b] by the Rocket Astronomy Group at the Naval Research Laboratory (NRL) have indicated that this radiation in the hard portion of the spectrum penetrates deeply into the *D* region and during large disturbances apparently produces ionization at altitudes of less than 60

km. The electron concentration is influenced not only by the intensity of the ionizing radiation but also by the recombination rate and by the ion detachment and attachment rates. It has been shown later, however, that the photoionization rate is sufficiently large so that the contribution of electron attachment to molecular oxygen is of negligible importance at altitudes above 60 km; this statement is true during the decay as well as during the build-up.

In section 2 we consider the NRL measurements of solar flare radiation and their implications with respect to the *D* layer, and on the basis of these data we construct a model of the ion production profile. In section 3 we estimate the dissociative recombination coefficient and construct a model of radio signal absorption based on the NRL data. Finally, in section 4 we

suggest an experiment for determining the recombination coefficients between 70 and 90 km.

2. *The nature and effects of ionizing radiation from solar flares.* Since World War II several theories have appeared regarding the nature of the radiation responsible for sudden ionospheric disturbances. The first was that the increase in ionization is caused by an enhancement in Lyman- α emission during a flare. The difficulties inherent in this suggestion are now apparent: Lyman- α would have to increase in intensity by a factor of $\sim 10^8$ or more in order to produce the high frequency absorption observed during sudden ionospheric disturbances. The detection of X rays in the rocket observations of solar flares (Sunflare I and II) by Friedman, Chubb, Kupperian, and Lindsay [1958]; Chubb, Friedman, and Kreplin [1960a] provides a basis for a much better explanation of the origin of sudden ionospheric disturbances. The X-ray flux is spread out in a spectrum which can cause sufficient ionization at all levels of the *D* region to produce the observed electron concentrations. It has also been suggested [Mitra, 1960] that both X rays and Lyman- α are important in producing SID's. However, recent satellite measurements of the Lyman- α intensity during a class 1⁺ solar flare [Chubb and others, 1960b] indicate that no such enhancement in the Lyman- α intensity occurs. A larger flare may, of course, produce an observable Lyman- α enhancement, but it is doubtful if it could compete with X radiation in its ionospheric effects at any level of the *D* region. All the measurements of the NRL group [Friedman, Chubb, Kupperian, and Lindsay, 1958; Chubb, Friedman, and Kreplin, 1960a, Chubb and others, 1960b] indicate that

the electromagnetic solar radiation producing the enhanced ionization in the *D* layer are X rays of wavelength $\lambda < 8 \text{ \AA}$; X rays of wavelengths $\lambda > 8 \text{ \AA}$ will be absorbed almost entirely in the *E* region. This is shown in Table 1,

TABLE 1. Intensity Required to Produce Given Electron Production Rates at Two Levels in the *D* Region

Altitude, km	Electron Production Rate q , $\text{cm}^{-3} \text{ sec}^{-1}$	Intensity, $\text{erg cm}^2/\text{sec}$		
		$2 > \lambda > 8 \text{ \AA}$	$\lambda > 8 \text{ \AA}$	$L - \alpha$
70 km	30	0.2	$\gg 1$	$\sim 10^6$
90 km	2500	0.2	~ 1	$\sim 10^5$

which gives the intensity at several wavelengths required to produce given electron production rates at two levels of the *D* region.

A partial X-ray flux spectrum was obtained from a Sunflare II rocket [Chubb, Friedman, and Kreplin, 1960a] launched during a class 2⁺ flare on August 31, 1959. Unfortunately, only a lower limit of the intensity in the 2- to 8- \AA range was observed owing to saturation of some of the sensors. Because of this gap in the spectral data it was necessary to estimate the missing portion by extrapolating from other data. Several other successful rocket flights in the Sunflare II program yielded essentially complete X-ray spectra in the range 2 to 60 \AA , but because they were made during a small flare and under quiet sun conditions, the cosmic-noise-absorption data were inadequate for the construction of a model of the disturbed *D* region. However, since we now had X-ray spectra for several lower levels of X-ray intensity, it was possible to extrapolate the shape of the X-ray spectrum to the level of X-ray intensity corresponding to the class 2⁺ flare considered in this paper. The spectrum so constructed, together with its estimated upper and lower limits (the broken curves) at time 22:54 UT, is shown in Figure 1. In addition, counter measurements in the energy range 15 to 80 keV were also made at two different times during the flight. These data indicated a softening of the hard X radiation as it decreased in intensity [Chubb, Friedman, and Kreplin, 1960a].

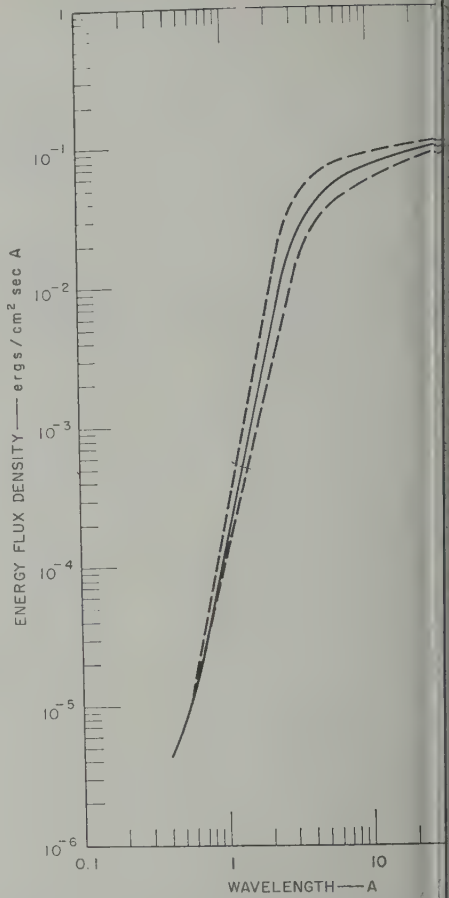


Fig. 1. X-ray spectrum of the flare at 22:50 UT on August 31, 1959. The solid line represents the most probable spectrum; the dashed lines represent extrapolated data. The broken lines represent estimated upper and lower limits of intensity.

When X rays are observed in the ionosphere they produce (a) primary ionization by photoelectric 'stripping' of electrons from atoms and molecules, and (b) secondary ionization by 'stripped' electrons. A photon of wavelength λ will produce about 200 free electrons per ionization process. The electron production rate q is written

$$q = \int \sigma_{\text{air}}(\lambda) \frac{dQ}{d\lambda}(\lambda) n \frac{E(\lambda)}{E_e} d\lambda$$

where $\sigma_{\text{air}}(\lambda)$ is the absorption cross section for X rays in air as a function of wavelength, $dQ/d\lambda$ is the X ray photon flux density

particle number density; $E(\lambda)$ is the energy of a photon of wavelength λ ; and E_0 is the mean energy required to remove an electron from a neutral molecule in air (~ 32 ev). Due to absorption by the atmosphere, the X-ray photon flux density $Q(\lambda)$ is also a function of altitude. Figure 2 shows the attenuation

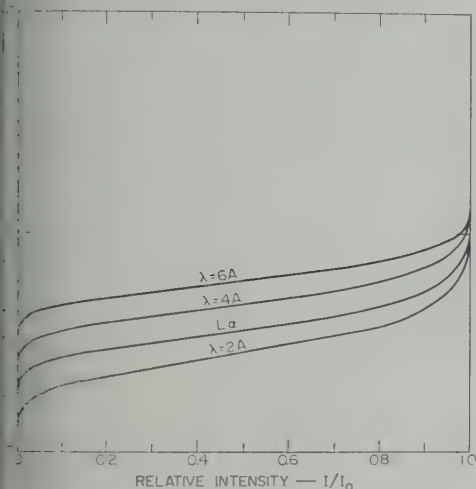


Fig. 2. Attenuation of solar radiation in the atmosphere. These curves are based on the 1959 model atmosphere [Minzner, 1959].

of α and X radiation of several wavelengths in the upper atmosphere. Figure 3 shows the electron-production-rate profile computed from the data presented in Figure 1.

We shall assume in our model that the electron-production rate bears the same functional relationship to time at all altitudes; this is equivalent to assuming that the shape of the X-ray spectrum is time-independent or that Q can be separated into two factors, one of which depends only on wavelength λ and altitude h , the other only on time

$$Q(\lambda, h, t) = \Lambda(\lambda, h)T(t) \quad (2)$$

We can establish the start of the X-ray emission at the time of the commencement of sudden cosmic-noise absorption (see Fig. 4) and the maximum of the X-ray emission curve at roughly the time of minimum cosmic-noise intensity. This is an oversimplification, since there is no reason to believe that the entire X-ray spectrum

bears exactly the same functional relationship to time.

3. *Model of ionization, recombination, and radio signal absorption.* A sudden ionospheric disturbance may be divided into two phases: the build-up phase and the decay phase. The former is characterized by a rapid increase in electron density due to the increasing solar X-ray photon flux and thus the electron production rate; the latter, on the other hand, is characterized by a gradual decrease in electron density. The electron-density maxima do not necessarily occur simultaneously at all altitudes, however, since there is at present no known reason for the intensity maxima of all parts of the X-ray spectrum to occur simultaneously. In spite of this, we shall assume in our model that the electron densities do simultaneously reach their maxima at all altitudes in the D region and that each maximum occurs at the same time as the minimum in the cosmic-noise-intensity curve (Fig. 4). The time rate of change of electron density

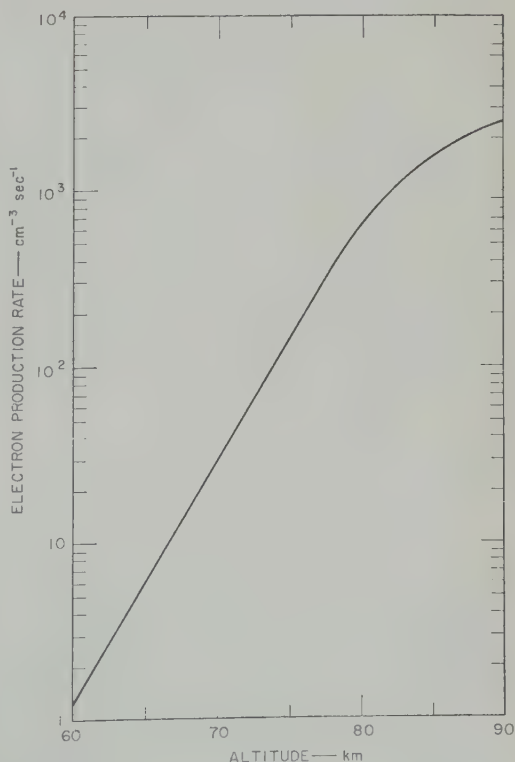


Fig. 3. Electron production rate at various altitudes as computed from the solid curve in Fig. 1.

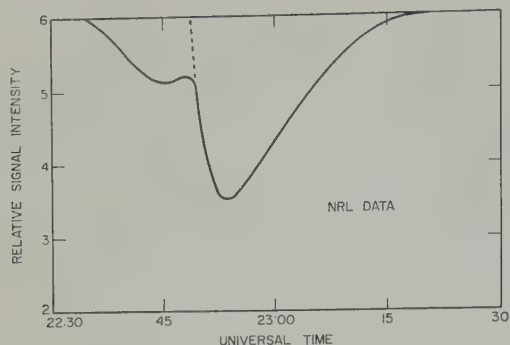


Fig. 4. Intensity at Los Angeles of 18 Mc cosmic noise during the occurrence of a class 2⁺ solar flare on August 31, 1959. The record indicates the extrapolation to its origin of the larger of the sudden cosmic-noise absorptions. Data reproduced through courtesy of U. S. Naval Research Laboratory.

is related to the production and removal processes by the well-known equation

$$\frac{dN_e}{dt} = \frac{q}{1 + \lambda} - (\alpha_D + \lambda\alpha_i)N_e^2 - \frac{N_e}{1 + \lambda} \frac{d\lambda}{dt} \quad (3)$$

in which N_e is the electron density; q is the electron production rate; λ is the negative ion to electron-density ratio ($\lambda = N^-/N_e$); α_D is the dissociative recombination coefficient; and α_i is the mutual ionic neutralization coefficient. The meaning of the symbol λ in the following (wavelength or negative ion to electron concentration ratio) will be clear from the context.

The differential equation obeyed by the negative ion to electron density ratio λ is [Bates and Massey, 1946]¹

$$\frac{1}{1 + \lambda} \frac{d\lambda}{dt} = \beta m - \lambda \left[\rho + \gamma n + N_e(\alpha_i - \alpha_D) + \frac{q}{N_e(1 + \lambda)} \right] \quad (4)$$

where β is the electron attachment coefficient; ρ is the photodetachment reciprocal lifetime; γ is the electron collisional detachment coefficient; n is the neutral particle density; and m is the molecular oxygen concentration (taken to be

1/5 n in our computations). In the model developed in this paper the neutral particle concentrations given in the ARDC 1959 atmosphere [Minzner, Champion, and 1959] are adopted.

The second term in brackets on the right hand side of equation 4 is of importance in the lower D region. Recent estimates of the collisional detachment coefficient γ based on thermal absorption [Bailey and Branscomb, 1960] are of the order $\gamma \approx 2 \times 10^{-17}$ cm³/sec; however, since other workers [Phelps and Phelps, 1961] have found evidence that $\gamma < 10^{-17}$ cm³/sec, the value of γ reduced by Bailey and Branscomb [1960] is probably at most a lower limit. Our investigation showed that the use of Bailey's value of γ only slightly affected the results (α_D , N_e (h , t), etc.) in the region below 90 km, thus indicating that collisional detachment is not very important in the solar disturbed D region. The third and fourth terms in brackets are, according to our calculations, approximately equal in magnitude and of opposite sign; hence they tend to cancel. The negative sign of the third term in [$N_e(\alpha_i - \alpha_D)$] is due to the fact that $\alpha_i < \alpha_D$ [Crain, 1961]. Because of its very slow dependence on N_e and q , λ is a slowly varying function of time² and can be approximated

$$\lambda \approx \frac{\beta m}{\rho + \gamma n}$$

Because of the weak time-dependence of λ , the last term on the right hand side of equation 4 can be neglected. The most recent investigations of electron-O₂ attachment [Chanin, Phelps, and Bondi 1959; Holt, 1959] indicate that in the D layer it is primarily a three-body process and that the coefficient K where

$$\beta = Km$$

has the value $K = (2 \pm 1) \times 10^{-30}$ cm³ at an electron temperature of 250°K. Equation 4a can then be written

$$\lambda \approx \frac{Km^2}{\rho + \gamma n}$$

¹ Mitra [1952, p. 314] presents equation 4 incorrectly in that in his work the electron source function q is not divided by $N_e(1 + \lambda)$.

² The insensitivity of λ to changes in N_e indicates that the approach to equilibrium attachment and detachment processes is rapid. This is merely the result of relative magnitudes of attachment and detachment coefficients.

photodetachment reciprocal lifetime ρ computed using recent measurements [Smith, and Branscomb 1958] of photodetachment cross section for O_2 and solar radiation measurements [Kuiper, 1953]. These give a value of $\rho = 0.35 \text{ sec}^{-1}$. The value of electron affinity of O_2 derived by Burch and Smith from the cross section data was 0.15 eV compared to the value of 0.46 eV obtained by Phelps and Pack [1961]. Since the photodetachment cross sections are expected to be increasing functions of the threshold energy of the process, our computed value of ρ may be large. Hence λ was also computed for values of ρ of 0.10 sec^{-1} and 0.035 sec^{-1} ; the λ values are shown in Figure 5.

If the values of K , ρ , γ , α_D , α_i , the neutral electron concentration profile, and the functional relationship to altitude and time of the electron production function $q(h, t)$ are known, it is possible to compute the electron concentration at any altitude h by means of the approximate form of equation 3:

$$\frac{N_e}{N_0} \approx \frac{q(h, t)}{1 + \lambda} - (\alpha_D + \lambda \alpha_i) N_e^2 \quad (3a)$$

Using the foregoing as a basis, we shall now construct a model of the D region during sudden ionospheric disturbances. To facilitate this we obtain an estimate of the dissociative recombination coefficient we shall utilize the NRL data and sudden cosmic-noise-absorption data presented in Figure 4. In essence our task is to find a functional form for q and α_D of the dissociative recombination coefficient α_D which, when inserted in equation 3a, will give electron concentration profiles at various altitudes during the build-up phase that agree with observed reduction in 18 Mc cosmic-noise intensity. The mutual ionic neutralization coefficient α_i does not enter the computations because of its small value relative to α_D [Crain, 1959] and because of the fact that the negative electron concentration ratio is at most of unity at altitudes of interest.

The relative cosmic-noise intensity is given by equation

$$I = I_0 \exp \left[-2 \int_D (k_d - k_n) dh \right] \quad (6)$$

where I_d and I_n are the 18 Mc cosmic noise

intensities corresponding to the disturbed and normal D region, respectively, and \int_D indicates integration over that segment of the ray path which lies in the D region. The absorption coefficient k for the case in which $\omega \gg \nu$ is given approximately by

$$k \approx \frac{5 \omega_0^2 \nu}{3 2 c \omega^2} \quad (7)$$

[Molmud, 1959], neglecting gyromagnetic splitting which is of little importance at 18 Mc; ω_0 is the plasma frequency, ω , the signal frequency, ν , the electron-neutral particle collision frequency, and c the velocity of light in free space. In constructing this model we used the electron collision frequencies obtained by Phelps and Pack [1959]. The X-ray spectrum given in Figure 1, together with equations 1 and 2, may be used to compute the electron-production-rate profile as a function of time if we assume a model for X-ray intensity time-dependence $T(t)$.

Several simple functional forms of $T(t)$, as well as several values of the dissociative recombination coefficient α_D , were used in equation 3a, which was then integrated numerically in order to obtain the electron concentration profiles at various times during the build-up phase. These

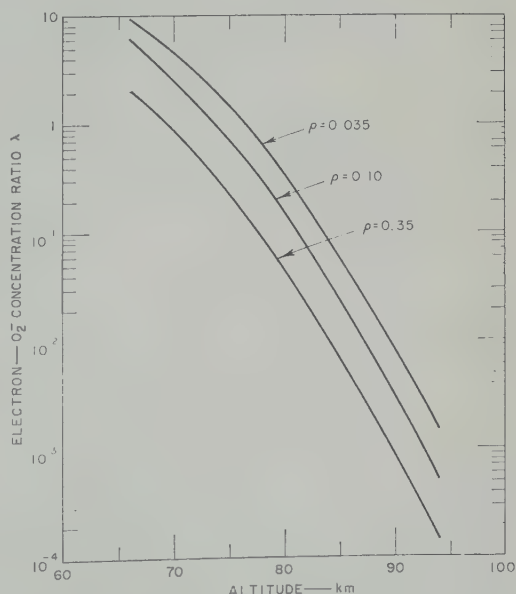


Fig. 5. Electron- O_2 concentration ratio profile for three possible values of the photodetachment coefficient ρ .

profiles were then used in conjunction with equations 6 and 7 to compute the relative cosmic-noise intensity I_a/I_n as a function of time. Upon comparison of $I_a(t)/I_n$ with the sudden cosmic-noise absorption data presented in Figure 1, it was immediately evident that only one of the trial analytical forms, i.e., $T(t) = a t^m$, was a reasonably valid choice. After various values of m in the range $\frac{1}{4} \leq m \leq 1$ were tried, it was found that $m = 1/2$ best agreed with the sudden cosmic noise absorption data. The value of α_D that best agreed with the SCNA data was $\alpha_D = (3 \pm 3) \times 10^{-7} \text{ cm}^3 \text{ sec}$; the uncertainty in α_D reflects the uncertainties in the X-ray spectrum and the electron collision frequency profile.

The decay phase was then treated in a similar manner. It was found that the functional form $T(t) = a'(t_0 - t)^2$ where $t_0 = 11.7$ minutes, and a value of $\alpha_D = 2 \times 10^{-7} \text{ cm}^3 \text{ sec}$ were, when used to compute the relative cosmic-noise intensity $I_a(t)/I_n$ in the manner outlined previously, in reasonable agreement with the sudden cosmic-noise-absorption observations until 15 minutes after commencement of the flare. The solid curve in Figure 6 shows the analytical form for $T(t)$ used in these computations.

The 18 Mc absorption exponent $2 \int_D (k_a - k_n) dh$ is shown as a function of time in Figure 7. The solid curve was obtained by substituting into equation 3a the analytical form of $T(t)$ and the values of α_D previously derived, then numerically integrating the equation to obtain the electron concentration profile, and substituting the electron concentrations $N_e(h, t)$

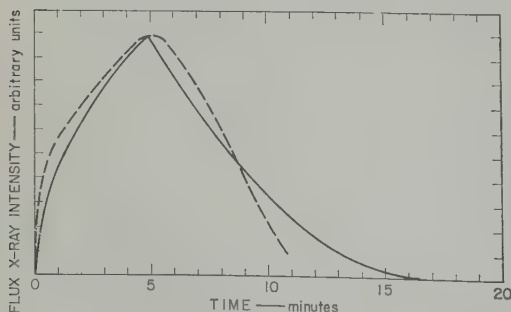


Fig. 6. Model of solar radiation intensity ($\lambda = 2$ to 8 \AA) as a function of time for the flare occurring on August 31, 1959. The solid line represents the function $at^{1/2}$, $0 \leq t \leq 5$ minutes and $a'(11.7 - t)^2$, $t > 5$ min. The broken line represents the estimated X-ray intensity.

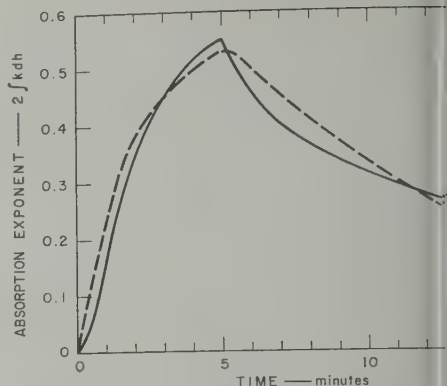
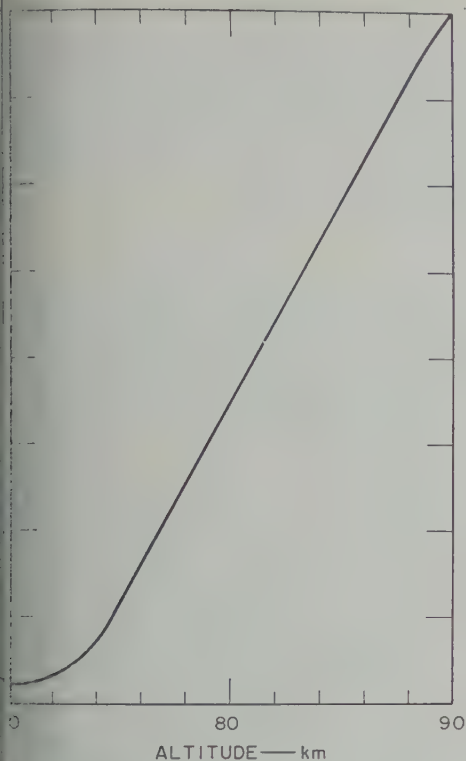


Fig. 7. The 18-Mc cosmic-noise-absorption exponent $2 \int k dh$ during the build-up and phases of the flare occurring at 22:50 UT on August 31, 1959. The broken curve represents the observed absorption and the solid curve represents the computed absorption using the function $(0 < t \leq 5 \text{ min})$ and $a(11.7 - t)^2$ ($5 < t < 11.7 \text{ min}$) to represent the electron ionization rate. The symbol k is equal to $k_a - k_n$.

so obtained into equation 7. The product $k_a(t) - k_n$ were then integrated over a path through the D region to obtain the absorption exponent. It is evident from Figure 7 the functional relationship to time of the intensity $T(t)$ actually increased faster than $at^{1/2}$ during the initial portion of the build-up phase and decreased faster than $a'(t_0 - t)^2$ during the decay phase. This is illustrated qualitatively by the broken curve in Figure 7. The broken curve agrees qualitatively with the intensity time-variation data obtained from different flares by satellite 1960 eta 22 (Kreplin, private communication). We reiterate that the factorization of $Q(\lambda, t)$ in accordance with equation 2 is quite satisfactory. It is known, for example, that the X-ray spectrum changes with time during the decay phase of a flare; this undoubtedly accounts for the disagreement between the observed and computed absorption exponents.

Figure 8 shows the electron concentration profile at 22:54 UT as computed by the method outlined previously. The corresponding absorption coefficient obtained by substituting the profile contained in Figure 8 into equation 7 is presented in Figure 9. This absorption coefficient profile is in rough agreement with the profile proposed by Mitra [1960].

Although it is self-consistent, the fit



Computed electron profile at time 22:54 UT.

is based on an assumption with respect to time variation of the flare spectrum that may not be justified. In order to determine its validity, a series of rocket probes launched during a sudden ionospheric disturbance and designed to measure electron density and flare spectral characteristics are necessary. Such an experiment is suggested in the following section. In the meantime, any flare spectral data that may be obtained from satellite observations, together with cosmic-noise observations, would be most useful in this respect.

4. proposed experiment. The model of ionospheric disturbances presented in the foregoing pages is based upon some assumptions concerning the solar flare X-ray spectrum and its time-dependence, and the electron collision frequencies. If one could somehow determine the evolution with time of the *D*-layer electron profile and the X-ray spectrum at the occurrence of the minimum in the cosmic-noise intensity curve, the model could be placed on a much more solid foundation. For this

purpose the following rocket probe experiment is suggested.

It is first necessary to choose a rocket capable of carrying the requisite instrument package to an altitude of greater than 100 km. For reasons that will become apparent later, the rocket must pass rapidly through the *D* layer. Hence the peak altitude must be well above 90 km. In addition, the vehicle must be reliable and easily launched. The last requirement is necessitated by the short time intervals between flights. The instrumentation to be installed in each package includes a probe for measuring electron density, three X-ray sensors for determining the X-ray spectrum, and an aspect sensor. These data are telemetered and recorded continuously. In order to obtain the data for constructing the model and for determining recombination coefficients, at least three vehicles should be launched during the decay phase of a sudden ionospheric disturbance at intervals of about 10 to 30 minutes, depending upon the magnitude of the associated solar flare. The occurrence of a suitable flare and an estimate of its magnitude can best be carried out by a flare patrol conducted by a near-by solar observatory. The launch time for the first rocket is most conveniently obtained from a record of cosmic noise intensity obtained at the launching site.

In analyzing the rocket data obtained during

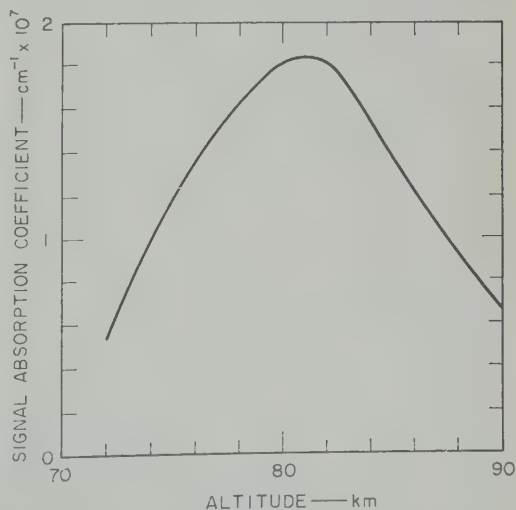


Fig. 9. The 18-Mc absorption coefficient profile corresponding to the electron concentration profile presented in Figure 7.

a sudden ionospheric disturbance, one must reduce the measured values of X-ray intensity to an electron-production-rate profile by means of equation 1. Since the values of Q are not measured simultaneously at the different altitudes, one must first standardize to a given time by estimating the time rate of change of Q at each altitude from the values observed at the same altitudes during the descent phase and during subsequent or previous flights. By numerical integration we obtain

$$\Delta Q = \int_{t_{\text{obs}}}^{t_{\text{stand}}} \left(\frac{dQ}{dt} \right) dt \quad (8)$$

and from this

$$Q_{\text{stand}} = Q_{\text{obs}} + \Delta Q \quad (9)$$

The value of q_{standard} is then obtained from equation 1.

The coefficients α_D and α_i are obtained by considering ionospheric relaxation at two altitude intervals: 80 to 90 km, and 60 to 75 km. In the 80- to 90-km interval the negative ion to electron concentration ratio λ is sufficiently small that one can neglect the term $\lambda\alpha_i N_e^2$ in comparison to α_D . Solution of the approximate differential equation

$$\frac{dN_e}{dt} \approx q(h, t) - \alpha_D N_e^2 \quad (10)$$

using an electron source function $q(h, t)$ derived from the X-ray flux data and initial and final N_e yields the coefficient α_D . The value of the mutual neutralization coefficient α_i (or in any event its upper limit) can be obtained from the data taken in the interval 60 to 75 km. In this interval the negative ion to electron concentration ratio λ becomes of order unity and greater. Computation of an 'effective recombination coefficient' $\alpha' = \alpha_D + \lambda \alpha_i$ at several altitudes in this range permits one to compute $\lambda\alpha_i$ by subtracting α_D from α' . In this range λ is given approximately by equation 4b which gives α_i to within approximately the uncertainty of ρ . In this respect further laboratory work on photo-detachment is desirable. Lack of a precise value of the collisional detachment coefficient also introduces some uncertainty. If α_i is less than about $10^{-8}/\text{cm}^2 \text{ sec}$, this approach will yield an upper limit only.

Acknowledgments. We wish to acknowledge the very generous provision of Sun Flare II data

in advance of publication by Dr. H. Friedman group at the Naval Research Laboratory as helpful discussions with Dr. George G. of the Air Force Research Division, Dr. Moreton of the Lockheed Solar Observatory, Dr. A. V. Phelps of Westinghouse Research Laboratories. This research was supported by the Advanced Research Projects Agency through contract AF 29(601)-199A.

REFERENCES

- Bailey, D. K., and L. M. Branscomb, Rat-
ficient for O_2 collisional detachment, *Br.
Phys. Soc.*, 5, 123, 1960.
- Bates, D. R., and H. S. W. Massey, The br-
actions in the upper atmosphere, *Proc. Roy.
A137*, 261-296, 1946.
- Burch, D. S., S. J. Smith and L. M. Bran-
Photo-detachment of O_2 , *Phys. Rev.*, 171,
175, 1958.
- Chanin, L. M., A. V. Phelps, and M. A.
Measurement of the attachment of slo-
trons in oxygen, *Phys. Rev. Letters*, 2, 3,
1959.
- Chubb, T. A., H. Friedman, and R. W. I.
Measurements made of high-energy X-ray
comparing three class 2+ solar flares,
Geophys. Research, 65, 1831-1832, 1960a.
- Chubb, T. A., and others, *IGY Bull. No. 423*,
AGU, 41, 717-721, 1960b.
- Crain, C. M., Ionization loss rates below 90-
Geophys. Research 66, 1117-1126, 1961.
- Friedman, H., T. A. Chubb, J. E. Kupper-
and J. C. Lindsay, X-ray emission of
flares, *IGY Rocket Rept.*, Ser. 1, 183-185,
1958.
- Holt, E. H., Electron loss processes in the
afterglow, *Bull. Am. Phys. Soc.*, 4, 112-113,
1959.
- Kuiper, G. P., editor, *The Sun*, University
Chicago Press, Chicago, 1953.
- Minzner, R. A., K. S. W. Champion, and
Pond, ARDC model atmosphere 1959, *Ad-
Surveys in Geophysics*, No. 115, Geo-
Research Directorate, Air Force Cambrid-
search Center, ARDC, August 1959.
- Mitra, A. P., Ionospheric effects of solar flares,
Ionization models, *Pennsylvania State
University Ionosphere Research Laboratory St-
Rept. No. 142*, Dec. 1, 1960.
- Mitra, S. K., *The Upper Atmosphere*, 2d ed.
Asiatic Society, Calcutta, 1952.
- Molmud, P., Langevin equation and the
ductivity of non-Maxwellian plasmas,
Rev. 114, 29-32, 1959.
- Phelps, A. V., and J. L. Pack, Electron
frequencies in nitrogen and in the lower
sphere, *Phys. Rev. Letters*, 3, 340-342,
1961.
- Phelps, A. V., and J. L. Pack, Collisional
ment in molecular oxygen, *Phys. Rev. Let-
111-113*, 1961.

(Manuscript received March 16, 1961;
June 19, 1961.)

Terrestrial Extremely Low Frequency Noise Spectrum in the Presence of Exponential Ionospheric Conductivity Profiles

JANIS GALEJS

*Applied Research Laboratory, Sylvania Electronic Systems
Waltham, Massachusetts*

Abstract. Calculations by Raemer of the terrestrial ELF noise spectrum which are based on a sharply bounded homogeneous ionosphere of frequency-dependent height and conductivity are extended to a nonhomogeneous ionosphere of exponential conductivity profile. The exponential ionosphere model provides slightly lower resonant frequencies and less damping of the resonant peaks than the sharply bounded ionosphere. The exponential model improves the agreement with noise spectrum measurements by Balser and Wagner.

Introduction. Observations of the extra-low frequency terrestrial noise spectrum show several marked peaks between 8 and 30 cps [Balser and Wagner, 1960]. Such peaks may be accounted for by considering the response of the earth-ionosphere cavity to lightning discharges [Schumann, 1957; Wait, 1960b and c]. Detailed calculations of the noise spectrum have been made by Raemer [1961b]. This analysis utilized lightning data compiled by Williams [1960] and a field expansion into a sum of resonant cavity modes [Schumann, 1957; Wait, 1960b and c] below a sharply bounded homogeneous ionosphere of frequency-dependent height (or) conductivity. Such an ionosphere was introduced in the ELF mode analysis by Pierce [1960] to account for measured attenuation in the frequency range 100 to 1000 cps. Raemer [1961b] assumes frequency-dependent ionosphere parameters to make the calculated resonance frequencies agree with the measured ones. This introduces high attenuation and the calculated resonant peaks are excessively damped.

An ionospheric model based on an exponential conductivity vs. height profile which accounts for measured ELF attenuation rates [Galejs, 1961] and provides a reasonable agreement with measured earth-ionosphere cavity resonance frequencies and Q factors [Balser and Wagner, 1960] has been developed recently [Galejs, 1961]. Noise spectra calculated on the basis of this exponential ionosphere may be expected to provide a better agreement with experiment than those calculated on the basis

of a sharply bounded homogeneous ionosphere, and will be investigated in this note.

The exponential ionosphere model is reviewed in section 2. In section 3 simple expressions are derived for the noise spectrum due to sources uniformly distributed over an increment of the polar angle. This spectrum representation is applicable to a homogeneous or a nonhomogeneous ionosphere provided that the propagation constant of the fields is known in the space between the earth and the ionosphere. Numerical calculations of the spectrum are discussed in section 4.

2. The exponential ionosphere. A model of the ionosphere based on an exponentially varying refractive index n has been introduced in the ELF wave analysis by Wait [1958, 1960a, b, and c]. Strictly speaking, his analysis is applicable only to an ionosphere that is sharply bounded at an altitude h where $|n| \gg 1$. Thus it only approximately accounts for the losses caused by lower ionosphere layers where this inequality is not satisfied.

Ionospheric conductivity profiles calculated from collision frequency and electron density data of Moler [1960], Nicolet and Aikin [1960], Waynick [1957], and the *Handbook of Geophysics* [1960] exhibit a nearly exponential height variation. The differential equations for wave propagation in such an ionosphere are soluble without imposing the restriction $|n| \gg 1$ [Galejs, 1961]. The propagation constant for the ELF waves is determined by an iteration process. An agreement within 1 cps with the measured resonant frequencies is obtained by a daytime

ionosphere approximated by

$$\sigma(z) = 10^{-6} \exp\left(\frac{z - 70}{3.25}\right) \quad (1)$$

(σ is in mhos per meter and z in kilometers) which is sharply bounded at a height h where

$$\sigma(h)/\omega\epsilon_0 = 1 = \text{constant} \quad (2)$$

This ionosphere model results in Q factors of 3.8 to 4.8 for the resonant modes in the frequency range between 10 and 30 cps. These Q values are slightly lower than the measured Q factors of 4 to 6 [Balser and Wagner, 1960].

The propagation constant $\gamma = ikS$ and the height h may be determined for $f < 30$ cps from the approximations

$$\begin{aligned} \text{Re } S &= 1.44 - 0.004f \\ -\text{Im } S &= 0.086 - 0.0003f \\ h &= 43 + 0.2f \quad (\text{km}) \end{aligned} \quad (3)$$

A more exact procedure should account for the different daytime and nighttime propagation characteristics of the respective hemispheres. However, the nighttime propagation constant was calculated with approximations that become inaccurate for frequencies of the earth-ionosphere cavity resonances [Galejs, 1961]. Such nighttime data will not be used in the present noise spectrum calculations.

3. *Noise spectrum of the spherical cavity.* The vertical electric fields of a vertical electric dipole of moment Id_s located at $\theta = 0$, $r = a$ may be represented [Wait, 1960b and c] by

$$E_r = \frac{I ds \eta}{4kha^2} \frac{\nu(\nu+1)}{\sin \nu\pi} P_\nu(-\cos \theta) \quad (4)$$

or

$$\begin{aligned} E_r &= \frac{I ds \nu(\nu+1)}{4\pi a^2 \epsilon_0 \omega h} \\ &\cdot \sum_{n=0}^{\infty} P_n(\cos \theta) \frac{2n+1}{n(n+1) - \nu(\nu+1)} \end{aligned} \quad (5)$$

where

$$k = \omega/c, \quad \eta = \sqrt{\mu_0/\epsilon_0} = 120\pi \text{ ohms},$$

a is the earth radius, h = height of the lower ionosphere edge, $P_\nu(-x)$ is a Legendre function, $P_n(x)$ is a Legendre polynomial, $\nu + \frac{1}{2} \approx kaS$,

and S is a root of the modal equation. These ionosphere cavity resonance frequencies are usually examined by study of equation 5 of Wait, 1957; Wait, 1960b and c; Raemer, 1960. It follows that E_r will exhibit maxima near

$$|n(n+1) - \nu(\nu+1)| = \min$$

or

$$n(n+1) \approx (kaS)^2$$

The same maxima occur near $|\sin \nu\pi| = \min$ in (4). This condition is equivalent to

$$\text{Re } \nu = ka \text{ Re } S - \frac{1}{2} = m$$

where m is an integer. Conversely, minima $|E_r|$ occur near $|\sin \nu\pi| = \max$ or

$$\text{Re } \nu = ka \text{ Re } S - \frac{1}{2} = m + \frac{1}{2}$$

The power received by a vertical stub antenna that is due to the above dipole will be

$$P = K |E_r|^2$$

where E_r is given by (4) or (5). The field representation (4) avoids a double summation in power calculations and will be used in the present analysis. Introducing the notation

$$K |I ds \eta|^2 = g(i\omega, \theta, \phi) a^2 \sin \theta d\theta d\phi$$

the received power spectrum becomes

$$\begin{aligned} G(i\omega) &= \int_0^{2\pi} \int_0^\pi g(i\omega, \theta, \phi) a^2 \sin \theta \\ &\cdot \left| \frac{\nu(\nu+1)}{4kha^2} \frac{P_\nu(-\cos \theta)}{\sin \nu\pi} \right|^2 d\theta d\phi \end{aligned}$$

where $g(i\omega, \theta, \phi)$ is proportional to the square of the dipole moment of the sources per unit volume. Assuming that

$$g(i\omega, \theta, \phi) = g(i\omega)$$

for $\theta_1 < \theta < \theta_2$ and is equal to zero elsewhere, (12) becomes

$$G(i\omega) = g(i\omega) \frac{2\pi\nu(\nu+1)\sigma(\sigma+1)}{16k^2 a^2 h^2 \sin \nu\pi \sin \sigma\pi}$$

$$\cdot \int_{\theta_1}^{\theta_2} P_\nu(-\cos \theta) P_\sigma(-\cos \theta) \sin \theta d\theta$$

where $\sigma = \nu^*$. This integral may be evaluated with the aid of (3.12.3) of Erdélyi [1953]. These relations are particularly simple for θ

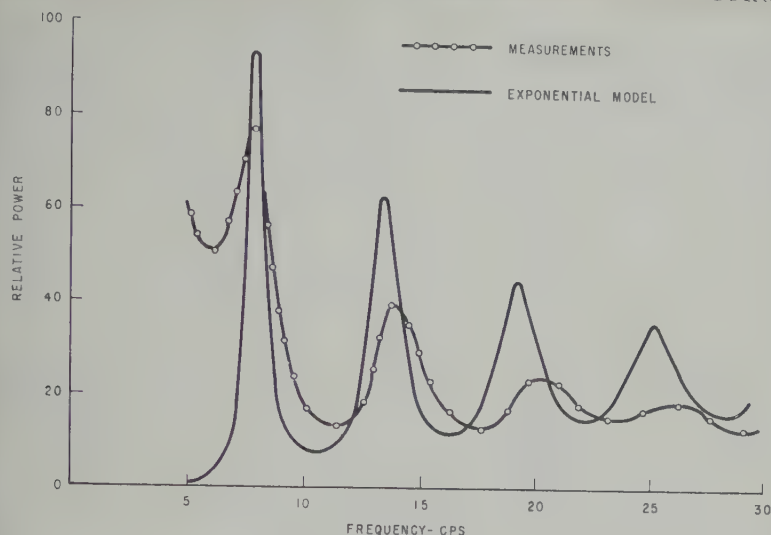


Fig. 1. ELF noise spectrum for a uniform distribution of white noise sources.

$= \pi$, when equations 3.12.7 and 1.7.2 of [1953] result in

$$\frac{g(i\omega)}{8h^2} \left| \frac{\nu(\nu+1)}{ak} \right|^2 \frac{1}{(-\text{Im } \nu)(\text{Re } \nu + 1/2)} \left\{ \frac{2}{\pi} \text{Im } \nu \sum_{n=0}^{\infty} \frac{1}{|\nu+1+n|^2} \frac{\sinh(2\pi \text{Im } \nu)}{\cos(2\pi \text{Re } \nu) - \cosh(2\pi \text{Im } \nu)} \right\} \quad (15)$$

integral (14) may also be evaluated after using the Legendre functions by their series expansions. It follows from 3.9.1.2 of [1953] that

$$\cos \theta \approx \sqrt{\frac{2}{\nu(\nu + \frac{1}{2}) \sin \theta}} \cos \left[\left(\nu + \frac{1}{2} \right) (\pi - \theta) - \frac{\pi}{4} \right] \quad (16)$$

results in

$$\frac{g(i\omega)}{8(akh)^2} \left| \frac{\nu(\nu+1)}{v + \frac{1}{2}} \right|^2 \left[\cosh(2\pi \text{Im } \nu) - \cos(2\pi \text{Re } \nu) \right] \left\{ \frac{\sinh[2 \text{Im } \nu(\pi - \theta)]}{(-\text{Im } \nu)} - \frac{\sin[(2 \text{Re } \nu + 1)(\pi - \theta) - (\pi/2)]}{\text{Re } \nu + \frac{1}{2}} \right\}_{\theta_1}^{\theta_2} \quad (17)$$

4. Numerical results. In the noise spectrum calculations the exponential ionosphere will be characterized by (3). The spectra are calculated first with $g(i\omega) = \text{constant}$ for a uniform source distribution in the polar angle θ from (15); they are shown in Figure 1. The resonant frequencies of the exponential ionosphere are slightly lower than the frequencies measured by Balser and Wagner [1960], which agrees with earlier calculations [Galejs, 1961]. The spectra calculated with $g(i\omega) = \text{constant}$ do not exhibit the higher-frequency decay observed in measurements.

On the basis of statistical data of Williams [1959], Raemer [1961a] has computed the power spectrum of the dipole moment in a median lightning flash. This curve (Fig. 2 of Raemer [1961b]) may be approximated by

$$g(i\omega) = 1.69 \exp[-9.1 \times 10^{-3} \omega] \quad (18)$$

Multiplying the spectra calculated in Figure 1 with (18) yields the curve of Figure 2, which provides a better agreement with the measurements.

In Figure 3 the noise spectrum calculated

with the exact integral (15) is compared with the asymptotic approximation (17). There is a qualitative agreement between the two curves which tends to justify the use of (17) in the noise spectrum calculations for nonuniform source distributions.

The effects of source distributions of equatorial widths is seen from Figures 1 and 2. The sources are assumed to be uniformly distributed over the polar angle θ in the range $(90^\circ - \Delta) \leq \theta \leq (90^\circ + \Delta)$. It is seen that a decrease of Δ increases the height of the main peak.

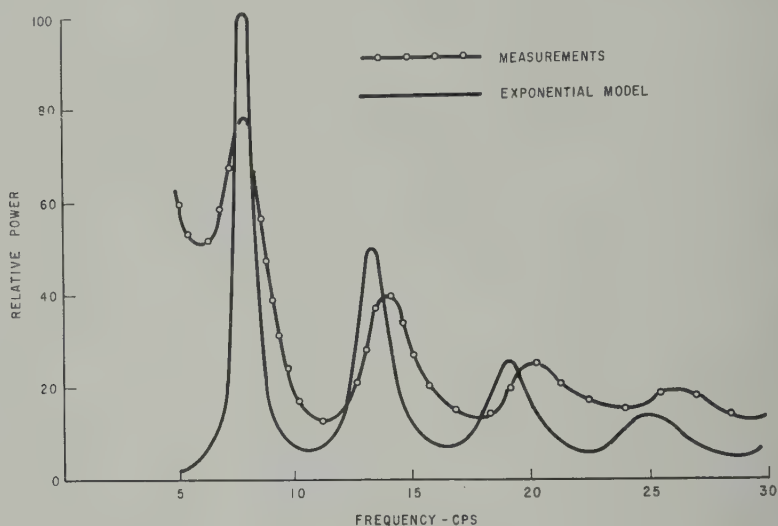


Fig. 2. ELF noise spectrum for a uniform source distribution. Spectrum of a median lightning flash assumed for all sources.

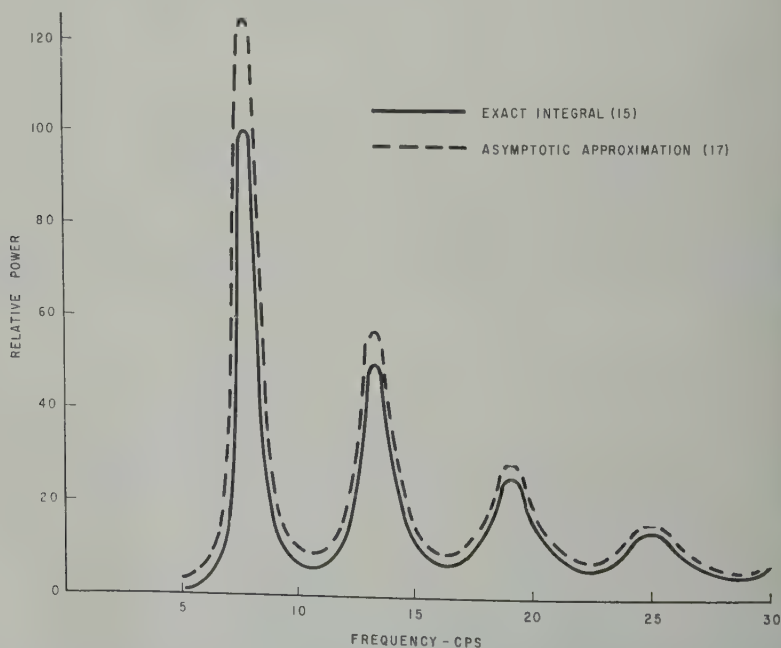


Fig. 3. ELF noise spectra based on the exact integral of the uniform source distribution and its asymptotic approximation. Spectrum of a median lightning flash assumed for all sources.

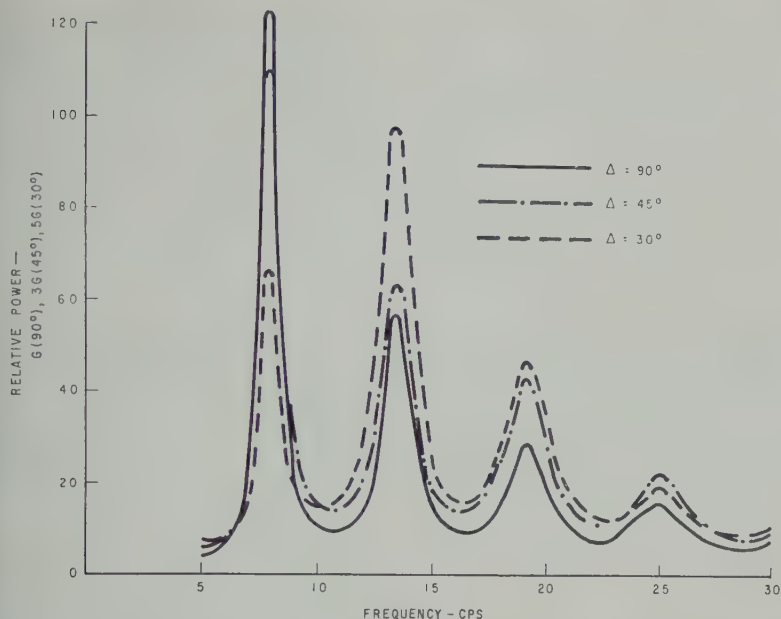


Fig. 4. ELF noise spectra $G(\Delta)$ for uniform source distribution over the polar angle interval $(90^\circ - \Delta) < \theta < (90^\circ + \Delta)$. Spectrum of a median lightning flash assumed for all sources. Asymptotic approximation of the integral.

of resonant peak G_2 relative to the first and peaks G_1 and G_3 . The observation point is related to a null of the electric field for odd values of Δ as Δ tends to zero. This is evident from pictures similar to those of Figure 6 of [1960b]. The calculated relative amplitudes of the first three spectral peaks are the same as the measured ones [Balser and Wagner, 1960], if $\Delta = 55^\circ$ to 60° in the exponential model. The principal regions of lightning activity occur for $\theta = 45^\circ$ to 135° with respect to Boston, Massachusetts. When assuming a uniform source distribution over this θ interval, Δ is about 45° , which is not too far from the value anticipated from Figure 5.

The exponential ionosphere model provides improved agreement with measured noise spectra, but several of the discrepancies require further clarification.

The relative amplitudes of the successive spectral peaks may be brought into agreement with measurements by defining a suitably wide uniform source distribution, but the peak-to-valley ratios of the exponential model are higher than the measured ones. The peak-to-valley

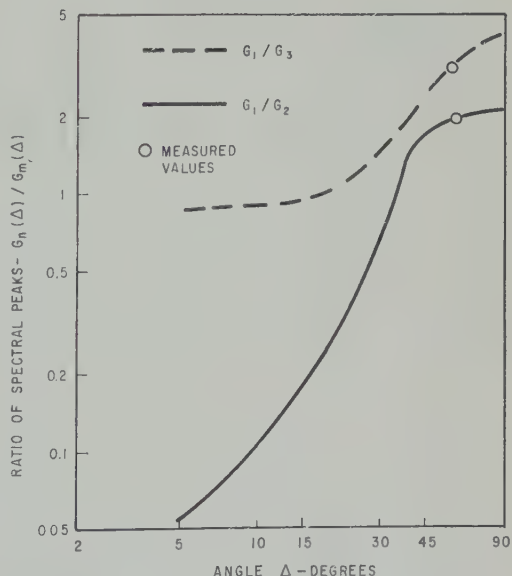


Fig. 5. Ratio of spectral peaks $G_n(\Delta)/G_m(\Delta)$ for a uniform source distribution over the polar angle interval $(90^\circ - \Delta) < \theta < (90^\circ + \Delta)$. Spectrum of a median lightning flash assumed for all sources. Asymptotic approximation of the integral.

ratios of the model may be decreased by decreasing of cavity Q or by introducing a constant level background noise. The suggested decrease of cavity Q is difficult to justify, since the calculated cavity Q is already lower than the measured Q . The constant level background noise is a more attractive speculation. It could be due to extraterrestrial noise or to terrestrial noise of local sources [Large and Wormsell, 1958; also Balser and Wagner, 1960].

More exact estimates of noise sources and an ionosphere model that accounts for the differences between day and night hemispheres merit further consideration, although the improvements possible by such refinements are difficult to assess. The magnetic field of the earth, which has been ignored in the present calculations, may influence nighttime propagation characteristics when the ELF fields reach higher altitudes where electron collision frequencies are lower.

Acknowledgment. This research has been supported by the Office of Naval Research. Project 79-300.

REFERENCES

- Balser, M., and C. A. Wagner, Observations of earth-ionosphere cavity resonances, *Nature*, 188, 638-641, 1960.
- Erdelyi, A., *Higher Transcendental Functions*, vol. 1, McGraw-Hill Book Co., New York, 1953.
- Galejs, Janis, ELF waves in the presence of exponential ionospheric conductivity profiles, *IRE Trans. on Antennas and Propagation*, AP-9, Nov. 1961.
- Handbook of Geophysics*, U. S. Air Force, ARDC, AFCRC, The Macmillan Co., New York, 1960.
- Jean, A. G., Jr., A. C. Murphy, J. R. Wait, and D. F. Wasmundt, Propagation attenuation rates at ELF, *J. Research NBS*, 65D, 1961.
- Large, M. I., and J. W. Wormsell, Fluctuations in the vertical electric field from 1 cycle per second to 500 cycles per second, *Recent Advances in Atmospheric Electricity*, edited by L. G. Galejs, Pergamon Press, New York, pp. 360-600, 1960.
- Moler, W. F., VLF propagation effects in the D -region layer produced by cosmic radio waves, *J. Geophys. Research*, 65, 1459-1468, 1960.
- Nicolet, M., and A. C. Aikin, The formation of the D region of the ionosphere, *J. Geophys. Research*, 65, 1469-1483, Figure 4, 1960.
- Pierce, E. T., The propagation of radio waves of frequency less than 1 kc, *Proc. IRE*, 48, 3, 1960.
- Raemer, H. R., On the spectrum of terrestrial radio noise at extremely low frequencies, *Technical Mem.*, Applied Research Lab., SRI, Electronic Systems, Waltham, Mass., Feb. 7, 1961a.
- Raemer, H. R., On the extremely low frequency spectrum of earth-ionosphere cavity resonance during electrical storms, *J. Geophys. Research*, 65, 1580-1583, 1961b.
- Schumann, W. O., Elektrische Eigenschwingungen des Hohlraumes Erde-Luft-Ionosphäre, *Z. Naturforsch.*, 9, 373-378, 1957.
- Wait, J. R., Extension to mode theory of ELF propagation, *J. Geophys. Research*, 63, 1, 1958.
- Wait, J. R., On the propagation of ELF waves and the influence of a nonhomogeneous ionosphere, *J. Geophys. Research*, 65, 1, 1960a.
- Wait, J. R., Terrestrial propagation of VLF waves—A theoretical investigation, *J. Research NBS*, 64D, 153-203, 1960b.
- Wait, J. R., Mode theory and the propagation of ELF radio waves, *J. Research NBS*, 64D, 404, 1960c.
- Waynick, A. H., Present state of knowledge concerning the lower ionosphere, *Proc. IRE*, 45, 741-749, June 1957.
- Williams, J. C., Thunderstorms and very low frequency noise, Ph.D. thesis, Division of Engineering and Applied Physics, Harvard University, June 1959.

(Manuscript received May 15, 1961)

Micrometeorite Measurements—Satellite and Ground-Level Data Compared

W. D. CROZIER

*New Mexico Institute of Mining and Technology
Socorro, New Mexico*

Abstract. Comparison is made between an annual mass accretion to the earth (3.9×10^5 metric tons) calculated on the basis of micrometeorite counts by satellites 1958 α and 1959 η , and an annual mass accretion of magnetic spherules (0.9×10^5 metric tons) determined by collections made at ground level by the Airborne Particle Study of the New Mexico Institute of Mining and Technology. Both measurements apply to the particle mass interval 3.3×10^{-10} g to 1.2×10^{-8} g. On the basis of the theory of E. J. Öpik, the ratio of the two amounts is quite reasonable, but the entire interpretation becomes uncertain because of the possibility that many of the particles encountered by the satellites may be in orbit around the earth.

Now and Alexander [1960], on the basis of micrometeorite counts made by satellites 1958 α (Explorer I) and 1959 η (Vanguard III), have given a tentative rate of mass accretion to the earth of approximately 10^5 tons per day, for a particle mass interval 1.2×10^{-10} g to 1.2×10^{-8} g. They also have given an indication of the mass accretion of the particles in this interval. The calculations are based upon measured impulse responses for the sensors on the satellites, combined with an assumed impact velocity of 30 km/sec.

The Airborne Particle Study of the New Mexico Institute of Mining and Technology, supported by the Office of Naval Research, for several years has conducted a program of continuous collecting and counting of magnetic spherules of presumed meteoritic origin. These spherules are deposited from the atmosphere by gravity and are exposed at isolated stations in central New Mexico. The trays contain a little glycol to catch particles settling into them, and the trays are given a magnetic separation as previously described [Crozier, 1960] to eliminate the bulk of ordinary dust. The magnetic spherules are mounted on Millipore disks, and the disk spherules are counted and classified under the microscope.

Magnetic separation is a necessity if an accurate study of spherule deposition from the atmosphere is to be made with a reasonable expenditure of effort. The elimination of nonmagnetic spherical particles of possible meteoritic

origin must be accepted, since it is almost impossible to make meaningful counts of such objects in the presence of the total fall of ordinary dust. A few nonmagnetic spherules have been seen in dust collections, but the present indications are that their number is not large compared with the number of magnetic spherules.

A question always is to be raised in regard to the industrial contribution to the magnetic spherule deposit. Such activities as fusion welding, grinding, smelting, and even the burning of coal, add to the atmosphere many spherical magnetic particles which are, in routine counting at least, indistinguishable from meteoritic spherules. At our collecting station 5 miles southwest of the small town of Magdalena, New Mexico, the industrial contribution should be comparatively low. Consideration of settling rates makes it appear unlikely that an appreciable number of spherules larger than 15- to 20-micron diameter will be carried from industrial areas to this site. The situation is not quite so clear for spherules in the 5- to 10-micron range and smaller, but it seems possible to infer from size distributions that the industrial contribution is small for these sizes also. The size distribution of magnetic spherules in the present atmospheric collections seems to be about the same as in collections made from sedimentary materials deposited long before the industrial era [Crozier, 1960].

According to theory developed by Öpik [1951, 1956], it is expected that an appreciable

fraction of micrometeorites arriving in the upper atmosphere—those within a certain range of combinations of velocity, particle size, and trajectory inclination—will be heated to the melting point but not vaporized, and then will descend to the surface as smooth spheres. The rate of infall for the spheres should therefore be smaller than the rate for all particles by a factor which, for the size interval 1.2×10^{-10} g to 1.2×10^{-8} g, probably should be between 2 and 20.

In addition to spherules melted in the atmosphere, it is conceivable that some of the micrometeorites may be in spherical form before entering the atmosphere. These could be a fraction of the particles resulting from collisions in the asteroidal belt [Whipple, 1955] and on the moon [Whipple, 1961]. Some of such particles, of small size and having a low velocity, penetrating the atmosphere without being melted, would contribute an unknown amount to the spherule deposit.

In Figure 1, the lower curve represents the cumulative size distribution in the rate of infall

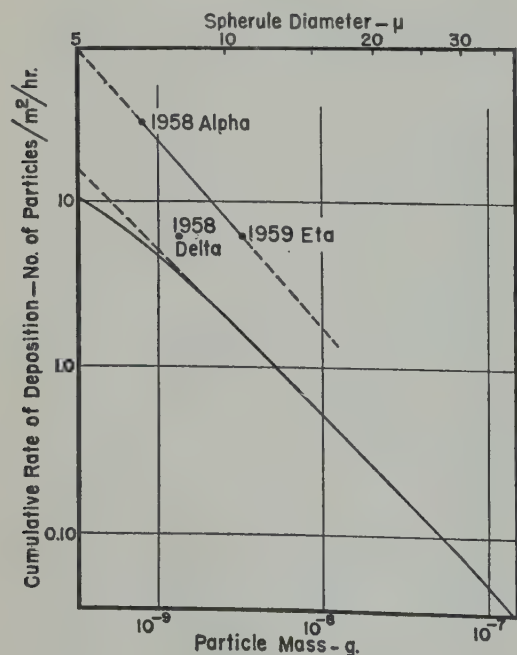


Fig. 1. *Upper Curve:* Cumulative rate of deposition of micrometeorites, calculated from satellite counts by LaGow and Alexander. *Lower curve:* Cumulative rate of deposition of magnetic spherules near Magdalena, N. Mex.

of magnetic spherules at the Magdalena. The counts include magnetic spherules deposited in a tray, usually with an area of 755 cm², during continuous exposure from April 1, 1960, to June 29, 1960. This is a much longer exposure than the one used previously [Crozier, 1960] to illustrate the size distribution, and the numbers are slightly different. The graph covers a range of spherule diameters from 5 to 30 microns. On the assumption that the density is 5 g cm^{-3} , the mass range is from 3.3×10^{-10} g to 1.5×10^{-7} g. The 5-micron lower limit is chosen because the recovery of spherules and the microscope work both become somewhat uncertain as the 5-micron size is approached. These data are used in the calculations and probably represent the true rate of deposition at this size as well as the full line. The distribution drops rapidly below the straight line as spherule size becomes larger than 35 microns in diameter. This is in qualitative agreement with the results of Thiel and Schmidt [1961] for the spherules in antarctic ice but is in contrast to the results of Laevastu and Mellis [1961] for spherules recovered from deep sea sediments.

The upper curve in Figure 1 includes the data points given by LaGow and Alexander [1960] at the thresholds for 1958 α and 1959 η . The moderate extrapolation in both directions and the extrapolation toward smaller sizes is carried to 3.3×10^{-10} g instead of 1.2×10^{-10} g as used by LaGow and Alexander in their calculations of mass accretion; all the present calculations are made to the lower limit of 3.3×10^{-10} g.

The apparent excellent parallelism between the two curves probably should be regarded as fortuitous. In fact, in the 38-micron interval, with velocities near 10^4 sec^{-1} , it would be expected that they would diverge toward larger masses, instead of showing the slight convergence of Figure 1.

Figure 1 also shows for comparison the threshold-impact rate point reported by Sarav, Nazarova, Neugodov, Poloskov, and Sarav [1958], for satellite 1958 δ , adjusted to an assumed impact velocity of 30 km sec^{-1} .

Using the two curves of Figure 1, the limits of 3.3×10^{-10} g and 1.2×10^{-7} g, the calculation of rate of mass accretion to the earth comes out 3.9×10^5 metric tons per year from the satellite data and 0.9×10^5 metric

for magnetic spherule deposition from the atmosphere. It should be emphasized strongly that neither of these rates represents the total rate of accretion; significant contributions are made by particles outside the limits in both dimensions. For example, integration carried out for the interval 3.3×10^{-10} g to 1.5×10^{-7} g on the magnetic spherule curve (lower in Fig. 1) gives a mass rate for this interval of 1.4×10^6 metric tons per year. Extrapolation of the satellite data to this interval would be unwise. If the two rates of accretion for the interval 1.5×10^{-7} g to 1.2×10^{-5} g are taken seriously, it would appear that about 23 per cent of the micrometeorites in this interval reach the earth in spherule form. On the basis of the Öpik [1956] theory, this is an entirely reasonable assumption; it is possible to construct a model of iron micrometeorites originally in angular form, having size distributions similar to those in Figure 1, and having velocities in the neighborhood of 30 km sec^{-1} , that will yield 23 per cent spherules.

However, we wish to proceed on the basis of the more elaborate Öpik [1956] theory, we must first modify the satellite curve to conform with Öpik's conclusion that the micrometeorites, if they are derived from the zodiacal cloud, enter the atmosphere with velocities between 11.1 and 12.2 km sec^{-1} , corresponding to unperturbed geocentric velocities in the range 5.1 to 5.5 km sec^{-1} . It is beyond the scope of this paper to carry out the calculations necessary for derivation of the rate of spherule accretion from the satellite data in this case, particularly since this calculation requires choice of a distribution of shape factors and of the ratio of mass to stony particles. Nevertheless, it seems probable that there is no inconsistency, at least in order of magnitude, between the satellite rate and the rate of spherule accretion.

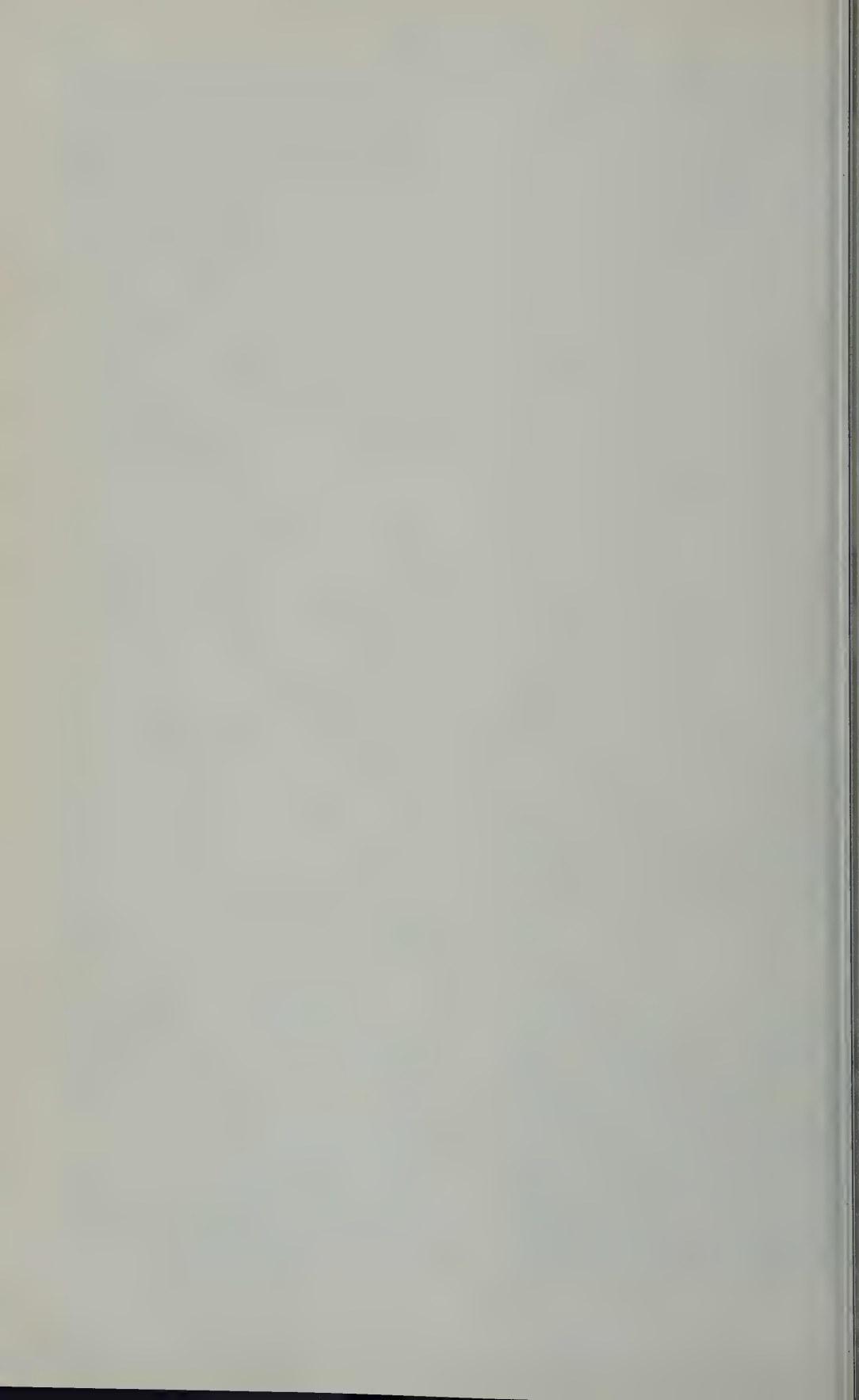
In the discussion thus far the possibility has been ignored that a large fraction of the micrometeorites intercepted by the satellites may have been in geocentric orbits. This possibility must be considered in the light of Hibbs' [1961] conclusion, based upon the apparent dependence of impact rate on altitude, that the 'average meteoroid' encountered by 1958 α was in a closed geocentric orbit. Two comments are necessary. Hibbs' conclusion is correct, the first that the data shown for the satellite points in Figure

1 must be increased by a factor of possibly 3 or 4, the second that practically none of the particles in geocentric orbits, at least in the 5- to 38-micron range, can be heated to the melting point on entering the atmosphere. However, it still is possible to account for spherule formation. Moving the satellite curve in Figure 1 to the right by a factor of 3 or 4 would leave space for a considerable nonorbiting component of the micrometeorite flux at the satellite. This component could be small enough to contribute only a minor fraction to the total flux while still being large enough to produce the observed rate of spherule accretion through operation of the Öpik melting process. For quantitative discussion of this kind of situation it seems to be necessary, however, to await further information on the micrometeorite population in space.

REFERENCES

- Crozier, W. D., Black, magnetic spherules in sediments, *J. Geophys. Research*, **65**, 2971-2977, 1960.
- Hibbs, A. R., The distribution of micrometeorites near the earth, *J. Geophys. Research*, **66**, 371-377, 1961.
- Komissarov, O. D., T. N. Nazarova, L. N. Neugodov, S. M. Poloskov, and L. Z. Rusakov, Micrometeorite studies using rockets and satellites, *Artificial Earth Satellites*, edited by L. V. Kurnasova, vol. 2, p. 68, Plenum Press, Inc., N. Y., 1960.
- Laevastu, T., and O. Mellis, Size and mass distribution of cosmic dust, *J. Geophys. Research*, **66**, 2507-2508, 1961.
- LaGow, H. E., and W. M. Alexander, Recent direct measurements of cosmic dust in the vicinity of the earth using satellites, *Proceedings of the First International Space Science Symposium*, 1033-1041, January 1960, North-Holland Publishing Company, Amsterdam. Also in *IGY Bulletin*, no. 38, p. 12, August 1960.
- Öpik, E. J., Astronomy and the bottom of the sea, *Irish Astron. J.*, **1**, 145-153, 1951.
- Öpik, E. J., Interplanetary dust and terrestrial accretion of meteoric matter, *Irish Astron. J.*, **4**, 84-135, 1956.
- Thiel, E., and R. A. Schmidt, Spherules from the antarctic ice cap, *J. Geophys. Research*, **66**, 307-310, 1961.
- Whipple, F. L., A comet model, *Astrophys. J.*, **121**, 750-770, 1955.
- Whipple, F. L., The dust cloud about the earth, *Nature*, **189**, 127-128, 1961.

(Manuscript received May 15, 1961; revised June 16, 1961.)



On the Long-Period Lunar and Solar Effects on the Motion of an Artificial Satellite, 2

PETER MUSEN

*Theoretical Division, Goddard Space Flight Center
National Aeronautics and Space Administration, Washington, D. C.*

Abstract. The disturbing function for the long-period lunisolar effects is developed into a series of polynomials in the components of the vectorial elements in the direction to the disturbing body. This development is convergent for all eccentricities and all inclinations. The equations are established for the variation of elements in a form suitable for the use of numerical integration and for the development of the perturbations into trigonometric series with numerical coefficients. An application of Milankovich's theory of perturbations leads to the equations for perturbed elements in which the small numerical divisors, the sine of the inclination and the eccentricity, are not present. These new equations, like the equations for canonical elements, have a symmetrical form and a wider range of applicability than the equations for elliptic elements.

SYMBOLS

M , the mean anomaly of the satellite for the moment t .

ω , the argument of the perigee of the satellite.

λ , the longitude of the ascending node of the satellite.

i , the inclination of the orbit of the satellite with respect to the equator.

e , the eccentricity of the orbit of the satellite.

a , the semimajor axis of the orbit of the satellite.

n , the mean motion of the satellite.

$(1 - e^2)$.

M' , the mean anomaly of the disturbing body at the moment t .

ω' , the argument of the perigee of the orbit of the disturbing body, but with respect to the ecliptic of the fixed epoch.

λ' , the longitude of the ascending node of the orbit of the disturbing body with respect to the ecliptic.

i' , the inclination of the orbit of the disturbing body toward ecliptic.

e' , the eccentricity of the orbit of the disturbing body.

a' , the semimajor axis of the orbit of the disturbing body.

n' , the mean motion of the disturbing body.

R , the disturbing function.

\mathbf{r} , the position vector of the satellite.

\mathbf{i} , the unit vector in the direction of \mathbf{r} .

\mathbf{P} , the unit vector directed from the center of the earth to the disturbing body.

\mathbf{j}, \mathbf{k} , the basic system of unit vectors in the equatorial system of coordinates.

\mathbf{P} , the unit vector directed from the center of the earth toward the perigee of the satellite.

\mathbf{R} , the unit vector standing normally to the orbit plane of the satellite.

\mathbf{Q} , the unit vector standing normally to \mathbf{P} and \mathbf{R} ; $\mathbf{Q} = \mathbf{R} \times \mathbf{P}$.

\mathbf{P}' , the unit vector directed from the center of the earth toward the perigee of the disturbing body.

\mathbf{R}' , the unit vector standing normally to the orbit plane of the disturbing body.

\mathbf{Q}' , a unit vector standing normally to \mathbf{P}' and

$$\mathbf{R}'; \mathbf{Q}' = \mathbf{R}' \times \mathbf{P}'.$$

α , the projection of \mathbf{P} on \mathbf{u}^0 .

β , the projection of \mathbf{Q} on \mathbf{u}^0 .

γ , the projection of \mathbf{R} on \mathbf{u}^0 .

$$\mathbf{c} = \sqrt{1 - e^2} \mathbf{R}.$$

$$\mathbf{g} = e\mathbf{P}.$$

$\nabla_a = \mathbf{i} \partial/\partial a_1 + \mathbf{j} \partial/\partial a_2 + \mathbf{k} \partial/\partial a_3$; the partial del operator with respect to vector

$$\mathbf{a} = ia_1 + ja_2 + ka_3.$$

Introduction. The object of this paper is the presentation of a theory of lunisolar effects on the motion of an artificial satellite in a form suitable for the use of the numerical integration procedure and for the development of the perturbations into Fourier series with numerical coefficients. The numerical integration procedure is less laborious than the development into

series, and the interval of integration can be of the order of several days. However, it is sometimes of interest to have the perturbations in the form of series, in particular, if a near-resonance effect occurs.

For satellites with a large ratio a/a' of semi-major axes the higher Legendre polynomials have to be included in the development of the disturbing function. The parallactic factor for a satellite with a period of 24 hours is about 0.1, and in considering the long-range stability it is advisable to take the parallactic term into consideration.

The problem of the development of perturbations into series can be solved either in an analytical or a semianalytical way. The purely analytical way is theoretically possible, but it presents considerable technical difficulties if the inclination of the orbit is not small.

Attempts made by *Musen, Bailie, and Upton* [1961] show that the technical difficulties become almost insurmountable with the inclusion of the fourth Legendre polynomial into the development of the disturbing function.

The use of a semianalytical procedure, valid for all inclinations and all eccentricities, is proposed in this paper. This method is adaptable to the use of electronic machines. The multiplica-

tion of Fourier series is performed automatically following the method developed at the NASA Aeronautics and Space Administration, Technical Division.

Development of the disturbing function. Notations in this article are the same as in the author's previous article [*Musen, 1961*] with the exception of the mean anomalies, which are designated in this paper by M and M' . The disturbing function Ω is taken in the form

$$\Omega = m'n \sum_{n=2}^{+\infty} p^{n+1} \left(\frac{a'}{r}\right)^{n+1} \cdot \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{r}{a}\right)^n P_n(\cos H) dM$$

p is the parallactic factor a/a' . Let \mathbf{r}^0 be a unit vector directed from the center of the earth to the satellite. We have

$$\mathbf{r}^0 = \mathbf{P} \cos f + \mathbf{Q} \sin f$$

and a corresponding unit vector for the sun (or the sun)

$$\mathbf{r}^0 = \mathbf{P}' \cos f' + \mathbf{Q}' \sin f'$$

where $(\mathbf{P}, \mathbf{Q}, \mathbf{R})$ and $(\mathbf{P}', \mathbf{Q}', \mathbf{R}')$ are the unit vectors connected with the osculating planes. We have in the equatorial system

$$\mathbf{P} = \begin{bmatrix} +\cos^2 \frac{i}{2} \cos(\omega + \Omega) + \sin^2 \frac{i}{2} \cos(\omega - \Omega) \\ +\cos^2 \frac{i}{2} \sin(\omega + \Omega) - \sin^2 \frac{i}{2} \sin(\omega - \Omega) \\ +\sin i \sin \omega \end{bmatrix}$$

$$\mathbf{Q} = \begin{bmatrix} -\cos^2 \frac{i}{2} \sin(\omega + \Omega) - \sin^2 \frac{i}{2} \sin(\omega - \Omega) \\ +\cos^2 \frac{i}{2} \cos(\omega + \Omega) - \sin^2 \frac{i}{2} \cos(\omega - \Omega) \\ +\sin i \cos \omega \end{bmatrix}$$

$$\mathbf{R} = \begin{bmatrix} +\sin i \sin \Omega \\ -\sin i \cos \Omega \\ +\cos i \end{bmatrix}$$

$$\left[\begin{aligned} & + \cos^2 \frac{i'}{2} \cos (\omega' + \Omega') + \sin^2 \frac{i'}{2} \cos (\omega' - \Omega') \\ & + \cos^2 \frac{i'}{2} \cos \epsilon \sin (\omega' + \Omega') - \sin^2 \frac{i'}{2} \cos \epsilon \sin (\omega' - \Omega') - \sin i' \sin \epsilon \sin \omega' \\ & + \cos^2 \frac{i'}{2} \sin \epsilon \sin (\omega' + \Omega') - \sin^2 \frac{i'}{2} \sin \epsilon \sin (\omega' - \Omega') + \sin i' \cos \epsilon \sin \omega' \end{aligned} \right] \quad (7)$$

$$\left[\begin{aligned} & - \cos^2 \frac{i'}{2} \sin (\omega' + \Omega') - \sin^2 \frac{i'}{2} \sin (\omega' - \Omega') \\ & + \cos^2 \frac{i'}{2} \cos \epsilon \cos (\omega' + \Omega') - \sin^2 \frac{i'}{2} \cos \epsilon \cos (\omega' - \Omega') - \sin i' \sin \epsilon \cos \omega' \\ & + \cos^2 \frac{i'}{2} \sin \epsilon \cos (\omega' + \Omega') - \sin^2 \frac{i'}{2} \sin \epsilon \cos (\omega' - \Omega') + \sin i' \cos \epsilon \cos \omega' \end{aligned} \right] \quad (8)$$

Elements of the satellite ω , Ω , i are referred to the equator, but the elements ω' , Ω' , i' of the disturbing body are referred to the ecliptic of a fixed epoch. Such a choice of reference permits us to consider i' as a constant and ω' and Ω' as linear functions of time, provided that we are interested in determining perturbations of the first order only.

We have from (2) and (3)

$$\mathbf{r} \cdot \mathbf{u}^0 = \alpha \cos f + \beta \sin f \quad (9)$$

$$\mathbf{r} \cdot \mathbf{P} \cdot \mathbf{u}^0 = \mathbf{P} \cdot \mathbf{P}' \cos f'$$

$$+ \mathbf{P} \cdot \mathbf{Q}' \sin f' \quad (10)$$

$$\beta = \mathbf{Q} \cdot \mathbf{u}^0 = \mathbf{Q} \cdot \mathbf{P}' \cos f'$$

$$+ \mathbf{Q} \cdot \mathbf{Q}' \sin f' \quad (11)$$

$$f' = M' + 2e' \sin M' + \dots \quad (12)$$

In addition, we put

$$\gamma = \mathbf{R} \cdot \mathbf{u}^0 = \mathbf{R} \cdot \mathbf{P}' \cos f' + \mathbf{R} \cdot \mathbf{Q}' \sin f' \quad (13)$$

We deduce from (4)–(8)

coeff.	α	β
$\cos^2 \frac{i'}{2} \sin^2 \frac{\epsilon}{2}$	$+\cos (\omega + \omega' + \Omega + \Omega' + f')$	$-\sin (\omega + \omega' + \Omega + \Omega' + f')$
$\cos^2 \frac{i'}{2} \cos^2 \frac{\epsilon}{2}$	$+\cos (\omega - \omega' + \Omega - \Omega' - f')$	$-\sin (\omega - \omega' + \Omega - \Omega' - f')$
$\cos^2 \frac{i'}{2} \cos^2 \frac{\epsilon}{2}$	$+\cos (\omega + \omega' - \Omega + \Omega' + f')$	$-\sin (\omega + \omega' - \Omega + \Omega' + f')$
$\cos^2 \frac{i'}{2} \sin^2 \frac{\epsilon}{2}$	$+\cos (\omega - \omega' - \Omega - \Omega' - f')$	$-\sin (\omega - \omega' - \Omega - \Omega' - f')$
$\sin^2 \frac{i'}{2} \cos^2 \frac{\epsilon}{2}$	$+\cos (\omega + \omega' + \Omega - \Omega' + f')$	$-\sin (\omega + \omega' + \Omega - \Omega' + f')$
$\sin^2 \frac{i'}{2} \sin^2 \frac{\epsilon}{2}$	$+\cos (\omega - \omega' + \Omega + \Omega' - f')$	$-\sin (\omega - \omega' + \Omega + \Omega' - f')$
$\sin^2 \frac{i'}{2} \sin^2 \frac{\epsilon}{2}$	$+\cos (\omega + \omega' - \Omega - \Omega' + f')$	$-\sin (\omega + \omega' - \Omega - \Omega' + f')$
$\sin^2 \frac{i'}{2} \cos^2 \frac{\epsilon}{2}$	$+\cos (\omega - \omega' - \Omega + \Omega' - f')$	$-\sin (\omega - \omega' - \Omega + \Omega' - f')$

coeff.	α	β
$\frac{1}{2} \cos^2 \frac{i}{2} \sin i' \sin \epsilon$	$+\cos(\omega + \omega' + \delta + f')$	$-\sin(\omega + \omega' + \delta + f')$
$\frac{1}{2} \cos^2 \frac{i}{2} \sin i' \sin \epsilon$	$-\cos(\omega - \omega' + \delta - f')$	$+\sin(\omega - \omega' + \delta - f')$
$\frac{1}{2} \sin^2 \frac{i}{2} \sin i' \sin \epsilon$	$-\cos(\omega + \omega' - \delta + f')$	$+\sin(\omega + \omega' - \delta + f')$
$\frac{1}{2} \sin^2 \frac{i}{2} \sin i' \sin \epsilon$	$+\cos(\omega - \omega' - \delta - f')$	$-\sin(\omega - \omega' - \delta - f')$
$\frac{1}{2} \sin i \cos^2 \frac{i'}{2} \sin \epsilon$	$-\cos(\omega + \omega' + \delta' + f')$	$+\sin(\omega + \omega' + \delta' + f')$
$\frac{1}{2} \sin i \cos^2 \frac{i'}{2} \sin \epsilon$	$+\cos(\omega - \omega' - \delta' - f')$	$-\sin(\omega - \omega' - \delta' - f')$
$\frac{1}{2} \sin i \sin^2 \frac{i'}{2} \sin \epsilon$	$+\cos(\omega + \omega' - \delta' + f')$	$-\sin(\omega + \omega' - \delta' + f')$
$\frac{1}{2} \sin i \sin^2 \frac{i'}{2} \sin \epsilon$	$-\cos(\omega - \omega' + \delta' - f')$	$+\sin(\omega - \omega' + \delta' - f')$
$\frac{1}{2} \sin i \sin i' \cos \epsilon$	$+\cos(\omega - \omega' - f')$	$-\sin(\omega - \omega' - f')$
$\frac{1}{2} \sin i \sin i' \cos \epsilon$	$-\cos(\omega + \omega' + f')$	$+\sin(\omega + \omega' + f')$

and

$$\begin{aligned}
 \gamma = & +\sin i \cos^2 \frac{i'}{2} \sin^2 \frac{\epsilon}{2} \sin(\delta + \omega' + \delta' + f') \\
 & +\sin i \cos^2 \frac{i'}{2} \cos^2 \frac{\epsilon}{2} \sin(\delta - \omega' - \delta' - f') \\
 & +\sin i \sin^2 \frac{i'}{2} \cos^2 \frac{\epsilon}{2} \sin(\delta + \omega' - \delta' + f') \\
 & +\sin i \sin^2 \frac{i'}{2} \sin^2 \frac{\epsilon}{2} \sin(\delta - \omega' + \delta' - f') \\
 & +\frac{1}{2} \sin i \sin i' \sin \epsilon \sin(\delta + \omega' + f') \\
 & -\frac{1}{2} \sin i \sin i' \sin \epsilon \sin(\delta - \omega' - f') \\
 & +\cos i \cos^2 \frac{i'}{2} \sin \epsilon \sin(\omega' + \delta' + f') \\
 & -\cos i \sin^2 \frac{i'}{2} \sin \epsilon \sin(\omega' - \delta' + f') \\
 & +\cos i \sin i' \cos \epsilon \sin(\omega' + f')
 \end{aligned}$$

expressions 14-16 for α , β , and γ are to be if the development of the perturbations series is required. If numerical integration is preferred, it is simpler to obtain α , β , and γ by computing independently expressions 4-6, 12, and 10-13.

Following an idea suggested by the author in a previous article [Musen, 1961], the disturbing function will be developed in powers of α and β and of using the standard procedure of expanding in powers of the inclination. This procedure leads to a more compact form of the disturbing function, and in the development into the numerical values of the inclinations are substituted into α , β , and γ from the table. We have from (9)

$$\cos H = \alpha \cos f + \text{sine term} \quad (17)$$

$$H = \frac{1}{2}(\alpha^2 + \beta^2) + \frac{1}{2}(\alpha^2 - \beta^2) \cos 2f + \text{sine terms} \quad (18)$$

$$H = \frac{3}{4}\alpha(\alpha^2 + \beta^2) \cos f + \frac{1}{4}(\alpha^3 - 3\alpha\beta^2) \cos 3f + \text{sine terms} \quad (19)$$

$$H = \frac{3}{8}(\alpha^2 + \beta^2)^2 + \frac{1}{2}(\alpha^4 - \beta^4) \cos 2f + \frac{1}{8}(\alpha^4 - 6\alpha^2\beta^2 + \beta^4) \cos 4f + \text{sine terms} \quad (20)$$

the cosine terms are retained, because the sine terms do not contribute any long-period terms in (1). Substituting 17-20 into 1, and using the following accurate expressions,

$$\int_0^{2\pi} \frac{r^2}{a^2} \cos 2f \, dM = +\frac{5}{2}e^2$$

$$\int_0^{2\pi} \frac{r^2}{a^2} \, dM = 1 + \frac{3}{2}e^2$$

$$\int_0^{2\pi} \frac{r^3}{a^3} \cos 3f \, dM = -\frac{35}{8}e^3$$

$$\int_0^{2\pi} \frac{r^3}{a^3} \cos f \, dM = -\frac{5}{2}e - \frac{15}{8}e^3$$

$$\int_0^{2\pi} \frac{r^4}{a^4} \cos 4f \, dM = +\frac{63}{8}e^4$$

$$\int_0^{2\pi} \frac{r^4}{a^4} \cos 2f \, dM = \frac{21}{4}e^2 + \frac{21}{8}e^4$$

$$\int_0^{2\pi} \frac{r^4}{a^4} \, dM = 1 + 5e^2 + \frac{15}{8}e^4$$

account, we deduce a compact form of the disturbing function convergent for all eccentricities and all inclinations.

We have

$$\begin{aligned} \Omega = m'np^3 \left(\frac{a'}{r} \right)^3 & \left[\left(+\frac{3}{4}\alpha^2 + \frac{3}{4}\beta^2 - \frac{1}{2} \right) \right. \\ & + (3\alpha^2 - \frac{3}{4}\beta^2 - \frac{3}{4}e^2) \\ & + m'np^4 \left(\frac{a'}{r} \right)^4 \left[\left(-\frac{75}{16}\alpha^3 - \frac{75}{16}\alpha\beta^2 + \frac{15}{4}\alpha \right) e \right. \\ & + \left(-\frac{25}{4}\alpha^3 + \frac{75}{16}\alpha\beta^2 + \frac{45}{16}\alpha \right) e^3 \Big] \\ & + m'np^5 \left(\frac{a'}{r} \right)^5 \left[\left(+\frac{105}{64}\alpha^4 + \frac{105}{32}\alpha^2\beta^2 \right. \right. \\ & + \frac{105}{64}\beta^4 - \frac{15}{8}\alpha^2 - \frac{15}{8}\beta^2 + \frac{3}{8} \Big) \\ & + \left(+\frac{315}{16}\alpha^4 + \frac{525}{32}\alpha^2\beta^2 - \frac{105}{32}\beta^4 \right. \end{aligned}$$

$$\begin{aligned} & - \frac{615}{32}\alpha^2 + \frac{15}{32}\beta^2 + \frac{15}{8}e^2 \Big) e^2 \\ & + \left(+\frac{105}{8}\alpha^4 - \frac{315}{16}\alpha^2\beta^2 + \frac{105}{64}\beta^4 \right. \\ & - \frac{135}{16}\alpha^2 + \frac{45}{32}\beta^2 + \frac{45}{64}e^4 \Big) e^4 \Big] + \dots \quad (21) \end{aligned}$$

A second concise form of the disturbing function can be obtained, if we eliminate α , β , and e from (21) in favor of vectors

$$\mathbf{c} = \sqrt{1 - e^2} \mathbf{R} \quad (22)$$

$$\mathbf{g} = e\mathbf{P} \quad (23)$$

These vectors are closely related to the constants of the area integral and to the constant of the Laplacian integral. Substituting

$$\alpha = \frac{\mathbf{g} \cdot \mathbf{u}^0}{e}, \quad \beta^2 = 1 - \alpha^2 - \gamma^2, \quad \gamma = \frac{\mathbf{c} \cdot \mathbf{u}^0}{\sqrt{1 - e^2}}$$

into (21), we deduce

$$\begin{aligned} \Omega = \frac{3}{4}m'np^3 \left(\frac{a'}{r} \right)^3 & [5(\mathbf{g} \cdot \mathbf{u}^0)^2 + (\mathbf{c} \times \mathbf{u}^0)^2 - g^2] \\ & + \frac{5}{16}m'np^4 \left(\frac{a'}{r} \right)^4 [-35(\mathbf{g} \cdot \mathbf{u}^0)^2 \\ & - 15(\mathbf{c} \times \mathbf{u}^0)^2 + 9g^2 + 12]\mathbf{g} \cdot \mathbf{u}^0 + \dots \quad (24) \end{aligned}$$

The expressions 21 and 24 are suitable for the application of numerical integration, and in a hidden form they contain all the long-period terms of the first order.

The variation of constants. Two systems of differential equations will be established. The first system will be for the classical elliptic elements, but the derivatives of Ω with respect to the elements will be eliminated in favor of derivatives with respect to α and β . This makes the actual computations easier. The second system of equations gives the vectors \mathbf{g} and \mathbf{c} in a direct way, and the derivation is based on the application of *Milankovich's* [1939] form of variation of constants. This theory is not widely used in celestial mechanics. Its mathematical apparatus belongs, possibly, more to the domain of theoretical physics than to the domain of classical celestial mechanics, but some recent investigations by *Popovich* [1960] show that the application of this theory to the secular perturbations of planets might be of great interest. The application of *Milankovich's* theory to the determination of the lunisolar effect leads to the equations that are free from the 'small divisors,' $\sin i$ or e , and consequently have a wider domain of validity than the equations for variations of elliptic elements.

We have from (4), (5), (10), and (11)

$$\frac{\partial \alpha}{\partial \omega} = +\beta$$

$$\frac{\partial \beta}{\partial \omega} = -\alpha$$

$$\frac{\partial \alpha}{\partial i} = +\gamma \sin \omega$$

$$\frac{\partial \beta}{\partial i} = +\gamma \cos \omega$$

and

$$\frac{\partial \alpha}{\partial \Omega} = \mathbf{u}^0 \cdot \mathbf{k} \times \mathbf{P} = \mathbf{u}^0 \cdot (\mathbf{k} \times (\mathbf{Q} \times \mathbf{R}))$$

$$= \mathbf{u}^0 \cdot (\mathbf{Q} \cos i - \mathbf{R} \sin i \cos \omega)$$

$$\frac{\partial \beta}{\partial \Omega} = \mathbf{u}^0 \cdot \mathbf{k} \times \mathbf{Q} = \mathbf{u}^0 \cdot (\mathbf{k} \times (\mathbf{R} \times \mathbf{P}))$$

$$= \mathbf{u}^0 \cdot (\mathbf{R} \sin i \sin \omega - \mathbf{P} \cos i)$$

or, in a final form,

$$\frac{\partial \alpha}{\partial \Omega} = \beta \cos i - \gamma \sin i \cos \omega$$

$$\frac{\partial \beta}{\partial \Omega} = \gamma \sin i \sin \omega - \alpha \cos i$$

Substituting

$$\frac{\partial \Omega}{\partial i} = \frac{\partial \Omega}{\partial \alpha} \gamma \sin \omega + \frac{\partial \Omega}{\partial \beta} \gamma \cos \omega$$

$$\frac{\partial \Omega}{\partial \omega} = \beta \frac{\partial \Omega}{\partial \alpha} - \alpha \frac{\partial \Omega}{\partial \beta}$$

$$\frac{\partial \Omega}{\partial \Omega} = \frac{\partial \Omega}{\partial \alpha} (\beta \cos i - \gamma \sin i \cos \omega)$$

$$+ \frac{\partial \Omega}{\partial \beta} (\gamma \sin i \sin \omega - \alpha \cos i)$$

into the equations

$$\frac{d\Delta M}{dt} = -\frac{1-e^2}{e} \frac{\partial \Omega}{\partial e} - 2a \frac{\partial \Omega}{\partial a}$$

$$\frac{di}{dt} = \frac{\cos i}{\sin i \sqrt{1-e^2}} \frac{\partial \Omega}{\partial \omega}$$

$$- \frac{1}{\sin i \sqrt{1-e^2}} \frac{\partial \Omega}{\partial \Omega}$$

$$\frac{d\Omega}{dt} = \frac{1}{\sin i \sqrt{1-e^2}} \frac{\partial \Omega}{\partial i}$$

$$\frac{d\omega}{dt} = + \frac{\sqrt{1-e^2}}{e} \frac{\partial \Omega}{\partial e} - \cos i \frac{d\Omega}{dt}$$

$$\frac{de}{dt} = - \frac{\sqrt{1-e^2}}{e} \frac{\partial \Omega}{\partial \omega}$$

we deduce for the long-period lunisolar

$$\frac{d\Delta M}{dt} = -\frac{1-e^2}{e} \frac{\partial \Omega}{\partial e} - 2p \frac{\partial \Omega}{\partial p} \left(\frac{1}{p} \right)$$

$$\frac{di}{dt} = \frac{\gamma}{\sqrt{1-e^2}}$$

$$\cdot \left(\frac{\partial \Omega}{\partial \alpha} \cos \omega - \frac{\partial \Omega}{\partial \beta} \sin \omega \right)$$

$$\frac{d\Omega}{dt} = \frac{\gamma}{\sin i \sqrt{1-e^2}}$$

$$\cdot \left(\frac{\partial \Omega}{\partial \alpha} \sin \omega + \frac{\partial \Omega}{\partial \beta} \cos \omega \right)$$

$$\frac{d\omega}{dt} = \frac{\sqrt{1-e^2}}{e} \frac{\partial \Omega}{\partial e} - \cos i \frac{d\Omega}{dt}$$

$$\frac{de}{dt} = - \frac{\sqrt{1-e^2}}{e} \left(\beta \frac{\partial \Omega}{\partial \alpha} - \alpha \frac{\partial \Omega}{\partial \beta} \right)$$

dition, we have

$$da/dt = 0$$

The development into series is required, we compute (26) in terms of ω , ω' , Ω , Ω' , f' using (16). In the process of computation several terms with the factors of the form $p^k \omega'/2$ can be neglected. The transition from M' is performed using the standard formulas therefore is not presented in detail. Because of the smallness of e' , only the first two or three terms of e' should be retained.

As an application and for a check we will solve the equations for variation of constants in the case of a close satellite. We put

$$m'n p^3 \left(\frac{a'}{r'} \right)^3 \left[\left(\frac{3}{4} \alpha^2 + \frac{3}{4} \beta^2 - \frac{1}{2} \right) + (3\alpha^2 - \frac{3}{4} \beta^2 - \frac{3}{4} e^2) \right]$$

Substituting this value into (26), we deduce

$$\begin{aligned} \frac{di}{dt} &= \frac{3m'n\gamma}{2\sqrt{1-e^2}} p^3 \left(\frac{a'}{r'} \right)^3 \\ &\quad [1 + 4e^2 \alpha \cos \omega - (1 - e^2) \beta \sin \omega] \\ \frac{d\Omega}{dt} &= \frac{3m'n\gamma}{2\sqrt{1-e^2}} p^3 \left(\frac{a'}{r'} \right)^3 \\ &\quad [1 + 4e^2 \alpha \sin \omega + (1 - e^2) \beta \cos \omega] \quad (27) \end{aligned}$$

$$\frac{d\omega}{dt} = \frac{3m'n}{2} p^3 \left(\frac{a'}{r'} \right)^3 (4\alpha^2 - \beta^2 - 1)$$

$$\sqrt{1-e^2} - \cos i \frac{d\Omega}{dt}$$

$$\frac{de}{dt} = -\frac{15m'ne}{2} p^3 \left(\frac{a'}{r'} \right)^3 \alpha \beta \sqrt{1-e^2}$$

The system is identical with the system deduced by the author in his previous article [Musen, 1961] using a less direct method.

The equations of Milankovich for the disturbed

with

$$\mathbf{c} = \mu \sqrt{a(1-e^2)} \mathbf{R}$$

$$\mathbf{g} = \mu^2 e \mathbf{P}$$

and the symbols ∇_c and ∇_g designate the partial del operators

$$\nabla_c = i \frac{\partial}{\partial c_1} + j \frac{\partial}{\partial c_2} + k \frac{\partial}{\partial c_3}$$

$$\nabla_g = i \frac{\partial}{\partial g_1} + j \frac{\partial}{\partial g_2} + k \frac{\partial}{\partial g_3}$$

For long-period lunisolar effects in the motion of an artificial satellite, the equations of Milankovich take a simpler form

$$\frac{d\mathbf{c}}{dt} = \mathbf{c} \times \nabla_c \Omega + \mathbf{g} \times \nabla_g \Omega \quad (28)$$

$$\frac{d\mathbf{g}}{dt} = \mathbf{g} \times \nabla_c \Omega + \mathbf{c} \times \nabla_g \Omega$$

with \mathbf{c} and \mathbf{g} defined by (22) and (23) and Ω defined by (24). The system (28) admits two particular integrals

$$\mathbf{c}^2 + \mathbf{g}^2 = 1$$

$$\mathbf{c} \cdot \mathbf{g} = 0$$

These integrals can be used either for a check or to reduce the computation to four components instead of six. As an application we give here the explicit form of the equations for the perturbations of a close satellite. It is not difficult to write these equations for more general cases. We have

$$\Omega = \frac{3}{4} m'n p^3 \left(\frac{a'}{r'} \right)^3 [5(\mathbf{g} \cdot \mathbf{u}^0)^2 + (\mathbf{c} \times \mathbf{u}^0)^2 - \mathbf{g}^2]$$

Taking

$$\nabla_g (\mathbf{g} \cdot \mathbf{u}^0)^2 = 2\mathbf{g} \cdot \mathbf{u}^0 \mathbf{u}^0$$

$$\nabla_g \mathbf{g}^2 = 2\mathbf{g}$$

$$\nabla_c (\mathbf{c} \times \mathbf{u}^0)^2 = \nabla_c (\mathbf{c}^2 - \mathbf{c} \cdot \mathbf{u}^0 \mathbf{u}^0 \cdot \mathbf{c})$$

$$= 2(\mathbf{c} - \mathbf{c} \cdot \mathbf{u}^0 \mathbf{u}^0)$$

$$\nabla_g \Omega = \frac{3m'n}{2} p^3 \left(\frac{a'}{r'} \right)^3 (5\mathbf{g} \cdot \mathbf{u}^0 \mathbf{u}^0 - \mathbf{g})$$

$$\nabla_c \Omega = \frac{3m'n}{2} p^3 \left(\frac{a'}{r'} \right)^3 (\mathbf{c} - \mathbf{c} \cdot \mathbf{u}^0 \mathbf{u}^0)$$

$$= \mathbf{c} \times \nabla_c R + \mathbf{g} \times \nabla_g R$$

$$= \mathbf{g} \times \nabla_c R + \frac{\mu^2 - \mathbf{g}^2}{c^2} \mathbf{c}$$

$$\times \nabla_g R - \frac{c^2}{g^2} \frac{\partial R}{\partial \tau} \mathbf{g}$$

into consideration, we deduce a system with no 'small divisors.'

$$\frac{d\mathbf{c}}{dt} = \frac{3}{2} m' n p^3 \left(\frac{a'}{r'} \right)^3 (5\mathbf{g} \times \mathbf{u}^0 \mathbf{u}^0 \cdot \mathbf{g} - \mathbf{c} \times \mathbf{u}^0 \mathbf{u}^0 \cdot \mathbf{c}) \quad (29)$$

$$\frac{d\mathbf{g}}{dt} = \frac{3}{2} m' n p^3 \left(\frac{a'}{r'} \right)^3 (2\mathbf{g} \times \mathbf{c} - \mathbf{g} \times \mathbf{u}^0 \mathbf{u}^0 \cdot \mathbf{c} + 5\mathbf{c} \times \mathbf{u}^0 \mathbf{u}^0 \cdot \mathbf{g}) \quad (30)$$

This form for equations is suitable for the determination of the influence of the sun on a close satellite. As is seen from (29) and (30), the change of \mathbf{c} is always in the plane normal to \mathbf{u}^0 and the change of \mathbf{g} can be decomposed into two components, the first one in the direction of \mathbf{Q} and the second normal to \mathbf{u}^0 .

In order to deduce the long-period perturbations caused by the moon, the process of averaging with respect to M' , the mean anomaly of the moon, must be applied to (29) and (30). We have

$$\begin{aligned} & \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{a'}{r'} \right)^3 \mathbf{u}^0 \mathbf{u}^0 \, dM' \\ &= \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{a'}{r'} \right)^3 (\mathbf{P}' \cos f' + \mathbf{Q}' \sin f') \cdot (\mathbf{P}' \cos f' + \mathbf{Q}' \sin f') \, dM' \\ &= \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{a'}{r'} \right)^3 \left\{ \frac{1}{2} (\mathbf{P}' \mathbf{P}' + \mathbf{Q}' \mathbf{Q}') \right. \\ & \quad \left. + \frac{1}{2} (\mathbf{P}' \mathbf{P}' - \mathbf{Q}' \mathbf{Q}') \cos 2f' \right\} \, dM' \end{aligned}$$

Taking

$$\mathbf{P}' \mathbf{P}' + \mathbf{Q}' \mathbf{Q}' = I - \mathbf{R}' \mathbf{R}'$$

where I is the idemfactor, and

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{a'}{r'} \right)^3 \, dM' &= (1 - e'^2)^{-3/2} \\ \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{a'}{r'} \right)^3 \cos 2f' \, dM' &= 0 \end{aligned}$$

into account, we have

$$\begin{aligned} & \frac{1}{2\pi} \int_0^{2\pi} \left(\frac{a'}{r'} \right)^3 \mathbf{u}^0 \mathbf{u}^0 \, dM' \\ &= \frac{1}{2} (1 - e'^2)^{-3/2} (I - \mathbf{R}' \mathbf{R}') \end{aligned}$$

Substituting this result into (29) and deduce for the long-period lunar effect

$$\frac{d\mathbf{c}}{dt} = \frac{3}{4} m' n p^3 (1 - e'^2)^{-3/2} (\mathbf{c} \times \mathbf{R}' \mathbf{R}' \cdot \mathbf{c} - 5\mathbf{g} \times \mathbf{R}' \mathbf{R}' \cdot \mathbf{g})$$

$$\frac{d\mathbf{g}}{dt} = \frac{3}{4} m' n p^3 (1 - e'^2)^{-3/2} [2\mathbf{c} \times \mathbf{g} + \mathbf{g} \times \mathbf{R}' \mathbf{R}' \cdot \mathbf{c} - 5\mathbf{c} \times \mathbf{R}' \mathbf{R}' \cdot \mathbf{g}]$$

By forming the dot product of (29') and deduce for the long-period lunar effect in eccentricity a formula analogous to the equation of (27)

$$\frac{de}{dt} = + \frac{15 m' n e \sqrt{1 - e'^2}}{4} p^3 \mathbf{P} \cdot \mathbf{R}' \mathbf{R}' \cdot \mathbf{Q} (1 - e'^2)$$

It is of interest to note that the long-period effect on close satellites depends only upon the position of the orbital plane of the moon and not influenced by the position of the perigee. In addition, the change of the velocity \mathbf{c} is always parallel to the orbital plane. The form of the perturbations is simple for more distant satellites, because terms depending upon the position of the perigee will appear in the parallactic term.

Conclusion. This article contains two developments of the averaged disturbance function to be used in the determination of long-period lunar and solar effects in the motion of an artificial satellite. The first form represents the development into series of polynomials with respect to cosines of the angles between the vectorial elements \mathbf{P} , \mathbf{Q} of the satellite and the position vector of the disturbing body. The second, more compact form utilizes the velocity and the Laplacian vector of the satellite to eliminate 'small divisors,' $\sin i$ and $\sin \omega$, from the equations for perturbations. The perturbations can either be developed into trigonometric series with numerical coefficients or they can be obtained by the method of numerical integration. The form of equations established here requires the interval of integration to be of the order of 10 days or even more.

The development of the perturbations into trigonometric series is of interest if the resonance condition occurs. The main terms

resonance conditions were listed by Cook [1]. Cook and Musen [1961] arrived quite independently at similar equations for lunisolar perturbations for close satellites. Cook's equations are similar to equations 27 of this article.

Acknowledgments. I should like to take this opportunity to express my gratitude to my colleagues, Mrs. Ann Bailie and Miss Maxine Goldsmith, for the careful preparation of the manuscript for print.

REFERENCES

Cook, G. E., Luni-solar perturbations of the orbit of an earth satellite, Report published by Royal Aircraft Establishment Farnborough, 1961.

- Milankovich, M., *Bull. Acad. Math. Natur. (A) Sci. Math. Phys.*, no. 6, Belgrade, 1939.
- Musen, P., On the long-period lunisolar effect in the motion of the artificial satellite, *J. Geophys. Research*, 66(6) 1659-1665, 1961.
- Musen, P., A. Bailie, and E. Upton, Development of the lunar and solar perturbations in the motion of an artificial satellite, *NASA Tech. Note, D-494*, 1961.
- Popovich, B., Secular perturbations of vectorial elements of planetary orbits, (Serb.), *Sci. Soc. Bosnia*, 14, ser. 3, pp. 1-50, 1960.

(Manuscript received June 8, 1961).



The Aeolosphere and Atmosphere of Venus¹

E. J. ÖPIK²

*Department of Physics, University of Maryland
College Park, Maryland*

Abstract. A self-consistent model of the Venus atmosphere that satisfies the multitude of existing observational data is proposed. Its main properties and the line of reasoning are described as follows. The radiative greenhouse effect cannot account for a surface temperature of 570°K. The blanketing must be due to dust; and the main source of energy, wind friction at the surface. The term 'aeolosphere' is proposed for the region between the surface and the clouds of Venus, where wind is responsible for grinding and raising the dust as well as for the heating. Self-consistency requires the opacity to be so high as to impose an adiabatic vertical lapse rate. The potential temperature in the aeolosphere must be nearly constant. Horizontal pressure gradients driving the wind are caused by climatic differences near the top of the dust clouds. The dust may consist of calcium and magnesium carbonates with impurities. Pressure estimates from CO₂ bands and optical scattering point to the presence of two distinct reflecting layers, the visible cloud level at about 0.6 atm and 340°K, and an upper haze level at 0.08 atm and 234°K. In the visual and near infrared the upper haze is virtually transparent, owing to forward scattering. In the ultraviolet, and in the infrared beyond 1.4 microns, it is opaque. Vertical cross sections of the Venus atmosphere are constructed for 80, 40, and 20 per cent CO₂. For 80 per cent CO₂ the results are somewhat more consistent, yielding a pressure of 4.3 atm at 570°K at the surface, 22 km below the visible cloud level. Despite uncertainties in the H₂O content, its amount is so small as to be nowhere condensable.

Introduction. As a basis of the present discussion, we assume the temperature around 570°K derived for the microwave range from 3 to 10 cm [Mayer, McCullough, and Sloanaker, 1960; Mayer, 1959] to be the temperature of the solid surface of Venus. The somewhat lower temperature at 0.8 cm [Gibson and McCullough, 1959] may be that of an atmospheric layer. An alternative interpretation of the microwave radiation by ionospheric emission leads to various difficulties [Sagan, Siegel, and Jones, 1960], and its origin as thermal radiation from the surface of Venus appears to be highly probable. Usually a greenhouse effect is suggested to maintain the high temperature; quantitative estimates of possible greenhouse shielding do not support this hypothesis, however. Instead, a high opacity of the atmosphere over the entire spectral range at thermal radiation must be assumed. It appears that radiation in both direc-

tions, toward and from the surface, is so completely impeded by the opacity of the Venus atmosphere (chiefly due to the particulate matter, perhaps dust from the surface), that a greenhouse effect cannot operate efficiently.

Instead of the minute or negligible amount of solar radiation transmitted by the dust clouds (the term 'dust' being used here to denote any kind of particulate matter), *wind friction* is suggested as a source of heating of the surface. On earth, aerodynamic friction dissipates about 2 per cent of the solar radiative input. It is shown here that a similar amount released at the surface of Venus may be amply sufficient to maintain, even necessarily lead to, the high temperature, if the opacity is high enough.

The term 'aeolosphere' is proposed accordingly, to denote the region on Venus between solid surface and visible clouds, where wind is responsible for the supply of heat as well as for grinding and raising the dust. Qualitatively, the dust-bowl model of Venus has been considered earlier [Öpik, 1950, 1956]; here a way of quantitative treatment is indicated.

The aeolosphere has no counterpart in the meteorological structure of the terrestrial at-

¹This paper was presented at the Forty-Second Annual Meeting, Sessions on Planetary Sciences, American Geophysical Union, Washington, D. C., December 21, 1961.

²On leave from Armagh Observatory, Northern Ireland.

mosphere. The analogue of the earth's surface, receiving most of the absorbed solar radiation, is the visible cloud level on Venus, situated 20 to 30 km above the solid surface. The region immediately above the visible cloud surface corresponds to the terrestrial troposphere; the heat absorbed at the cloud surface is transmitted upward by convection and radiation. The troposphere is followed upward by a Venus stratosphere, of a stable thermal gradient with purely radiative exchange (Fig. 1).

2. *Greenhouse effect by molecular absorption.* From a comprehensive analysis of all available data [Öpik, 1961] (cf. sections 5 and 8), we distinguish on Venus three effective radiating or reflecting layers: layer R in the stratosphere, responsible for the observed thermal radiation of effective temperature $T_e = 234^\circ\text{K}$ [Sinton and Strong, 1960]; layer C of the visible cloud

surface, which is the main reflector and absorber of solar radiation, of a mean temperature $T_c = 340^\circ\text{K}$; and layer S, or the solid surface, at $T_s = 570^\circ\text{K}$. Between C and R the exchange is mainly radiative, the opacity being predominantly due to particulate matter and only partly to molecular absorption (CO_2); this region is accessible to observation and less problematic. Here we are concerned primarily with the energy exchange in the aeolosphere, between the solid surface and visible clouds, or layers S and C.

In purely radiative exchange, the greenhouse equation can be written schematically as

$$t_i J = t_e Q$$

where J = total absorbed solar radiation, t_i = fraction of J reaching the solid surface, Q = unshielded emission from surface to space, t_e = fraction of Q actually lost by the surface.

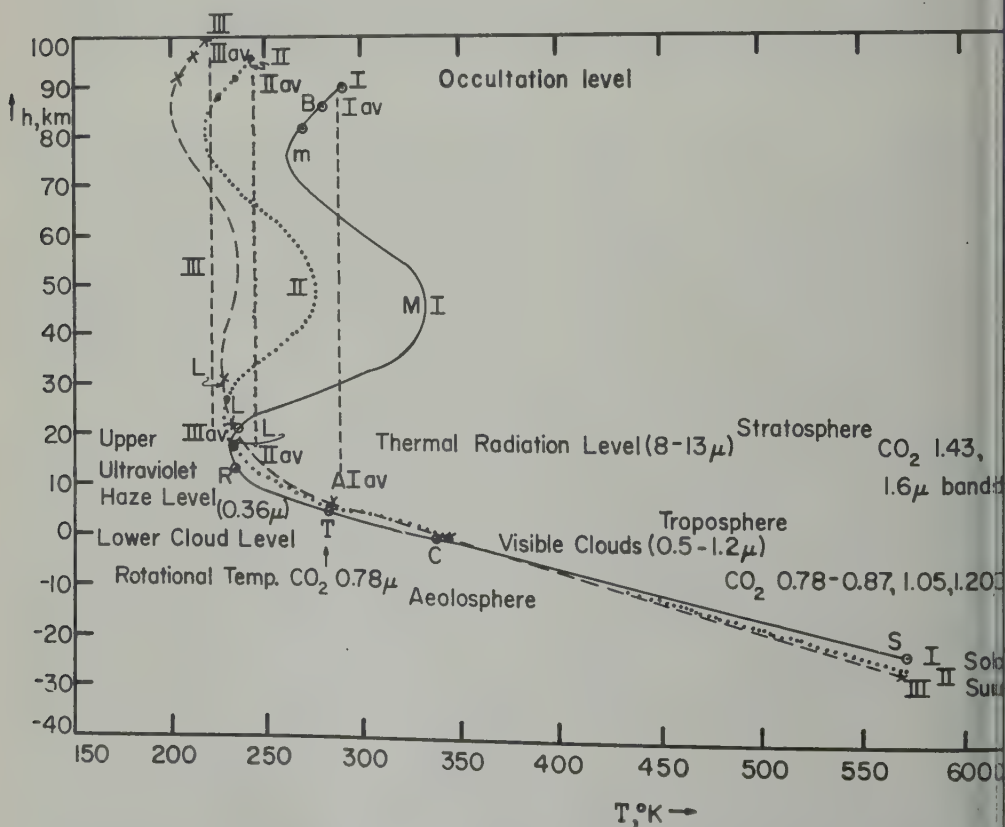


Fig. 1. Vertical cross section of the atmosphere of Venus. Abscissas, temperature $^\circ\text{K}$; ordinates, altitude km above visible cloud level. I for 80 per cent, II for 40 per cent, III for 20 per cent CO_2 by weight.

numerical estimates, we assume the black-body expression

$$Q = s(T_s^4 - T_c^4) \quad (2)$$

where $s = 5.67 \times 10^{-5}$ erg/cm² sec deg⁴. Also,

$$J = sT_c^4 \quad (3)$$

By definition, the absorbed radiation equaling that emitted to space. The cloud surface is assumed to be opaque for all radiation in the thermal range.

From equations 1, 2, and 3 the pure greenhouse condition for our model is obtained as

$$t_i/t_e = (T_s^4 - T_c^4)/T_c^4 \quad (4)$$

Substituting the numerical values for the temperatures, we find

$$t_i/t_e = 31 \quad (5)$$

the condition of efficient radiative greenhouse effect in the aeolosphere.

Only two molecular absorbents are definitely known to exist on Venus, CO₂ and H₂O.

Water vapor, in the amount derived by Sagan (private communication, 1960), 0.002 g/cm² above the visible clouds, can play but an insignificant role. Its total amount above the cloud surface, as of a noncondensable constituent conserving a constant mixing ratio, is then of the order of 0.04 g/cm² along the average optical ray path, implying a transmission of about 70 per cent of thermal radiation [cf. Lythe, 1956, p. 546]. Attempts to upgrade the water-vapor content of the Venus atmosphere cannot be accepted; in the 1.13-μ region the reflected light comes from the deep clouds (see C or T, cf. section 8 and Table 3), and the pressure is not much below terrestrial atmospheric, so that a pressure correction to Sagan's value cannot be significant. Weakly justified as it is, a suggestion by Sagan [1961, footnote], to introduce a correction to Sagan's value for pressure broadening as appropriate to the so-called square-root region of the rate of growth, may be considered; the equivalent width Δλ of a band is then represented by

$$\Delta\lambda = C(pm)^{1/2} \quad (6)$$

where p = pressure, m = absorbing mass per unit centimeter column, and C is a proper constant for a given band. For given Δλ, m is inversely proportional to p , and a smaller pres-

sure would mean more H₂O on Venus. As is shown in section 9, the water-vapor content can be increased only 3 to 6 times by this extreme assumption, and the transmission of thermal radiation at 570°K will still be of the order of 50 per cent.

Actually, for the observed H₂O band no pressure correction seems to be required, and, whatever the possible error in Strong's determination, the water-vapor content of the Venus atmosphere must be extremely low (cf. also section 9).

Carbon dioxide is a more important absorber on Venus, owing to its abundance. However, wide portions of its spectrum are free from absorption. Exaggerating the absorption, by assuming it to be complete in the band regions [List, 1951; Landolt-Börnstein, 1951, p. 330], we find that radiation will escape from the Venus surface practically unhindered in the wavelength intervals of 3-4, 5.5-9, 11-12.5, and >18 microns. At $T_s = 570^\circ\text{K}$ nearly two-thirds of the blackbody radiation will be transmitted. Combined with water vapor, the lower limit of transmission is then

$$t_e > 0.5$$

The maximum possible value of t_i is 1. Hence, for a greenhouse effect through molecular absorption,

$$t_i/t_e < 2$$

which is utterly inadequate (cf. equation 5). With these extreme assumptions, including the improbable case of all the nonreflected solar radiation penetrating to the surface, from equation 4 the upper limit of surface temperature which can be sustained by the molecular greenhouse effect results as 373°K, or +100°C.

The gap between requirements and fulfillment is so wide that no unidentified gaseous constituents could essentially alter the picture. Even when there is absorption over all the continuous spectrum, according to the theory of radiative transfer the effective opacity k_e is defined as the harmonic mean weighted according to spectral intensity, $E(\lambda)$ (Planck's curve),

$$k_e^{-1} = \int [k(\lambda)]^{-1} E(\lambda) d\lambda / \int E(\lambda) d\lambda \quad (7)$$

This causes strong bias in favor of low opacities, so that the more transparent regions of the

spectrum gain greatly in importance. Hence, with the uneven absorption by molecular bands, radiation will leak through the spectral 'windows,' greatly reducing the blanketing by the greenhouse.

Terrestrial water vapor, whose absorption bands cover nearly 75 per cent of the thermal spectrum (at 2 g/cm² precipitable water), causes a greenhouse effect of only $T_s/T_e = 288/250 = 1.15$, $t_i/t_e = 1.75$, and hardly much more can be expected on Venus from molecular constituents.

3. *Blanketing by dust.* The model of molecular absorption alone must be rejected as being unable to produce the required blanketing on Venus, yet surface radiation is efficiently prevented from reaching the cloud level and leaking out to space. The obvious conclusion is that particulate matter ('dust') must be responsible for the screening. It is but natural to assume that the dust that constitutes the visible clouds [Öpik, 1956] extends down to the very surface, possibly getting thicker and of coarser grain with greater depth.

Despite unlimited screening ability, dust can lead to a strong greenhouse effect only under very particular conditions. Although it is known from the theory of radiative transfer that omnidirectional scattering and true absorption may equally contribute to screening of a radiative flux, the scattering power of particles of a given size decreases with increasing wavelength. Hence, as far as scattering is concerned, the ratio t_i/t_e will tend to be less than unity, a circumstance not conducive to a greenhouse effect.

One favorable case is that of the incoming solar radiation being weakly scattered and absorbed, the outgoing thermal radiation being strongly absorbed by the dust. For this, the diameter $2r$ of the particles must be considerably less than λ/π (Rayleigh scattering), or, with $\lambda = 0.5$ micron as the median wavelength of solar radiation, less than about 0.1 micron. Polarization data, however, suggest particle diameters in the range of 2.5 microns (water droplets), or even 5 to 10 microns (quartz) [Van de Hulst, 1952] for the visible clouds. In such a case, solar radiation will be efficiently scattered by refraction and reflection, if not by diffraction, as is actually demonstrated by the high reflecting power of the planet. Apparently,

the case of very small particles does not to Venus.

With the diameters of the particles in the range of 2 to 5 microns or greater, and increasing with depth, the effective cross section for omnidirectional scattering of sunlight by transparent droplets is of the order of 0.01 according to the calculations of Wiener [Van de Hulst, 1952, Table 5], as depending on refraction and reflection. In addition, diffraction scattering corresponds to the geometrical cross section, πr^2 , but for large particles this is predominantly forward scattering, which does not affect the radiative flux and can therefore be neglected. For the outgoing thermal radiation in the wavelength interval of 4 to 20 microns, if the particles are of the optimum size for absorption, reflection, refraction, and diffraction combined can lead to a maximum cross section for omnidirectional scattering of $2\pi r^2$. The optical depth, τ , being proportional to the cross section, from equation 10 below, for large values of τ , the extreme possible greenhouse ratio comes

$$t_i/t_e \leq 2\pi r^2 / (0.25\pi r^2) = 8$$

which is far short of the condition set by equation 5.

From equation 4, the upper limit of the surface temperature that can be maintained by this type of greenhouse results as

$$T_s < 439^\circ \text{K}$$

The conditions here assumed are too extreme to be plausible. There is some absorption by the visual clouds, as is indicated by their optical depth of less than 1, making the effective cross section about $0.32\pi r^2$ for sunlight. On the way out, the particles will scatter thermal radiation, not only in a forward direction, and absorption can be complete, so that an effective cross section about πr^2 is more likely. Hence $t_i/t_e \sim 3$. $T_s \sim 390^\circ \text{K}$ is a more acceptable estimate of greenhouse screening by the dust. It seems that neither of the two mechanisms taken separately or together holds any promise of explaining the high surface temperature by a greenhouse effect.

4. *Aeolosphere.* The preceding section arrives at two main conclusions: (a) the atmosphere must be filled with dust, and (b)

ative greenhouse effect is inadequate to maintain the high surface temperature.

to supply the outward heat flux from the hot surface, however small it may be, a source of energy other than solar radiation must be available.

Wind friction at the earth's surface dissipates an equivalent of 1 to 2 per cent of the solar radiative input, two orders of magnitude less than supplied by conductivity from the interior. A similar source can be assumed to operate at the surface of Venus. Zonal and meridional differences in temperature at the cloud-level must lead to differences in pressure which drive atmospheric circulation reaching down to the aeolosphere and causing dissipation of energy by wind friction.

On a dry land, at an average wind velocity of about 400 cm/sec near ground level and a coefficient of aerodynamic friction of about $\eta = 0.01$, the work of friction

$$q = \eta \rho w^3 \quad (8)$$

is of the order of 4×10^3 ergs/cm² sec, with $\rho = 1.3 \times 10^{-3}$ g/cm³ as the density of air. On the earth, with $\eta = 0.01$ and $w = 800$ cm/sec, $q = 10^3$, or about the same. The average input of solar radiation is 2.5×10^5 ergs/sec/cm² of the earth's surface, whence friction amounts to about 2.4 per cent of the input.

In the aeolosphere of Venus is a passive region where the circulation is induced as a secondary effect by unequal heating and pressure differences, Δp , at the cloud top (level C). For given wind speed and coefficient of friction, the relation

$$\rho w^2 = \text{constant}$$

approximately holds, whence the work of friction according to equation 8, varies as $\rho^{-1/2}$. At the surface of Venus, where the atmospheric density is about ten times that at level C, wind speed may be about three times smaller than at the surface of a solid surface extending to the cloud level C itself. On the other hand, at the rate of the planet's rotation, Coriolis deflection must be less important, and wind will more directly follow the barometric gradient and may attain higher speeds. For sample calculations, we may thus assume that 2 per cent of the radiative input is dissipated by friction at the surface, the heat of friction being

$$q = \alpha J \quad (9)$$

with $\alpha = 0.02$; the coefficient, however, could be several times smaller.

From the theory of radiative transfer, the screening factor equals

$$i_c = (1 + 1.5\tau)^{-1} \quad (10)$$

where τ is the optical depth. Equations 10, 1, and 2 define the flux equally for omnidirectional scattering and for absorption and reradiation. In the first case the temperature distribution is irrelevant; when τ is not wavelength dependent, the intensity of original radiation is decreased without alteration of its relative spectral distribution. In the second case, a proper radiative equilibrium distribution of temperature is established in the space between the two levels S and C.

Equation 10 applies to the flux transmitted in multiple scattering and reradiation: by definition of optical depth, the fraction of the original radiation transmitted directly equals

$$I/I_0 = \exp(-\tau \sec z) \quad (11)$$

where z is the angle between line of sight and radial (vertical) direction.

Substituting equation 9 for $t_s J$, and t_s from equation 10 in equation 1, with equations 2 and 3 the radiative transfer condition for the dust-blanketed aeolosphere becomes

$$1 + 1.5\tau = (T_s^4 - T_c^4)/(\alpha T_c^4) \quad (12)$$

For $\alpha \leq 0.02$ and the adopted temperatures,

$$\tau \geq 10^3 \quad (13)$$

is obtained as a lower limit.

On the other hand, the dust content of the aeolosphere is determined by independent factors, and so is its optical depth.

At a flux as defined by equation 9, in radiative equilibrium the temperature gradient is proportional to the opacity or, more or less, to τ ; when the opacity is small, the gradient is stable and pure radiative equilibrium holds. When opacity and τ increase, the gradient increases until the adiabatic value is reached when turbulence sets in. Convective transport takes care of the part of the flux that cannot be transported by radiation; this requires but a very small excess of the gradient over the adiabatic value [Öpik, 1953], so that, at arbitrarily large τ values, any given flux can be accommodated.

Hence two basically different types of aeolospheric structures are to be considered.

One has an adiabatic or very slightly super-adiabatic vertical temperature gradient, convectionally unstable and turbulent down to the surface. The presence of dust as carried upward by turbulence in this atmosphere can be readily understood, the top of the dust clouds making the upper boundary of turbulence. This condition requires a high initial dust content, whose opacity stirs up turbulence, which, in its turn, raises the dust. The cycle may thus be self-supporting.

The other type has a small initial dust content, small opacity, and a stable thermal gradient. Convection in the aeolosphere then does not operate, whereas horizontal turbulent flow cannot carry the dust to any appreciable altitude. The dust sinks and sticks to the surface, the opacity decreases further, and the condition is again self-supporting and stable. Besides, with an appreciably stable thermal gradient, planetary atmospheric circulation will tend to die down in the aeolosphere, in the absence of energy sources there. The region would tend to become stagnant, with little motion, so that the frictional source of energy would also disappear. Of course, a downward transport of heat by forced turbulence of horizontal flow [Öpik, 1953] would become active, but only in the upper part of the region where the horizontal motion has not died down. The transported heat, proportional to the square root of the gradient deviation from the adiabatic value, asymptotically vanishes when the adiabatic gradient is approached; therefore, it cannot upset the stable gradient, and the once-established condition cannot be changed by it. The presence of thick dust clouds cannot be reconciled with this type of atmospheric structure.

Thus, with the acceptance of the model of a dust-laden atmosphere, there is no other choice but to accept also, as a necessary condition, the turbulent structure of the aeolosphere, with a near-adiabatic vertical lapse rate. Further, the dust must be real dust, nonvolatile and non-sublimating; condensable particulate matter would form cloud sheets with clear spaces in between, like terrestrial clouds; these could not build up an efficient protective blanket to shield radiation from the surface. Magnesium and calcium carbonates, mixed with silicate material,

could be suggested as the parent material of the dust.

The average particle diameter of these should decrease with altitude, for obvious reasons. In the upper, ultraviolet-absorbing layer it may be around 0.5 micron (cf. section 7), at the main visible cloud level around microns, and perhaps 5 to 10 microns in the aeolosphere. With an effective cross section to πr^2 with respect to combined absorption and scattering, the optical depth equals the geometrical coverage

$$\tau_0 = \pi r^2 n$$

where n is the number of particles per square centimeter of vertical column. With $r = 10^{-4}$ cm and equation 13 this gives

$$n \geq 5 \times 10^9$$

particles per square centimeter or, for 25 km the depth of the aeolosphere, an average of 200 dust particles per cubic centimeter or equivalent compact thickness of the dust

$$d = 4r\tau/3$$

is then of the order of 0.3 cm or more, which might be sufficient to cause some absorption of microwave radiation at 8.6-mm wavelength, offering a plausible alternative explanation to pure molecular absorption as suggested by Barrett [1960].

As to sunlight, for pure scattering without absorption a not negligible fraction of it as defined by equation 10 could reach the surface. But pure scattering from a layer of practically infinite optical depth would lead to an albedo of unity, which is not the case. With an albedo of 0.6 to 0.75 as for the Venus clouds, the effective scattering is of the order of $a = 0.8$ [Barrett, 1960], so that a fraction $1 - a = 0.2$ is absorbed in each scattering process. This makes the fraction of sunlight penetrating to a solid surface of Venus equal to

$$\exp(-0.2\tau) = \exp(-200)$$

or nil. With respect to sunlight, the aeolosphere of Venus is immersed in complete darkness.

As to microwave radiation, in the absence of particles larger than 0.1 cm, it cannot be efficiently scattered, only absorbed. The scattering by particles with

$$2\pi r < \lambda$$

re λ = wavelength, is omnidirectional, but the scattering cross section, decreases rapidly with wavelength,

$$\sigma \cong \pi r^2 (2\pi r / \lambda)^4 \quad (16)$$

(Rayleigh scattering). This explains why micrometeorites can pass through the aeolosphere without obstruction except when there is true absorption.

No attempt is made at this stage to offer a model of atmospheric circulation in the aeolosphere. It may only be pointed out that two conditions must be met by such a model. The first stresses the constancy of temperature at a given level; the heat capacity of the aeolosphere, $c = 5000 \text{ g/cm}^2$ and $c = 8 \times 10^6 \text{ ergs/g deg}$, and the order of $4 \times 10^{10} \text{ ergs/cm}^2 \text{ deg}$, and the differences in wind friction are of the order of $0.2q$ or $10^8 \text{ ergs/cm}^2 \text{ sec}$; hence possible rates of temperature change, in an atmospheric current moving through the climatic belts, are of the order of $10^3/4 \times 10^{10} \text{ deg/sec}$, less than a degree per year. For a circulation time of 10 to 20 days, the average temperature of the aeolospheric air could only vary by an insignificant amount of 0.02 to 0.04 deg. The second condition is that of the adiabatic lapse rate. The combination of the two requires a simultaneous adiabatic distribution of temperature to apply everywhere through the aeolosphere, so that temperature is a unique function of altitude or that the temperature is constant. Climatic and thermal differences in temperature can occur only at the cloud level where sunlight is absorbed. As to horizontal pressure gradients in the aeolosphere, they would lead to temperature gradients from adiabatic expansion were it not for wind friction, which converts the major part of the work of expansion into heat, so that little adiabatic cooling can take place, and the temperature will remain almost constant at the given altitude.

Structure of the atmosphere of Venus. A detailed study of the surface properties of Venus has been made by Öpik [1961]. In an attempt to satisfy all the available observational evidence, including photometric, polarimetric, spectroscopic, thermal radiometric, and micro-meteoritic data, a model has been constructed as a true, internally consistent solution. New evidence led to some modifications, without basic changes, in a previous model [Öpik, 1956]. Un-

doubtedly further modifications may be needed, but the chances appear to be good for the survival of the basic features of the model, it being fitted to a score of independent limiting conditions that could not be expected to agree accidentally.

A list of the main lines of evidence used in setting up the model may illustrate this point: (1) spectral-energy distribution and relative albedo from ultraviolet to red; (2) brightness distribution over the disk in ultraviolet and red; (3) brightness contrast of bands in ultraviolet and blue; (4) polarization in visual and ultraviolet; (5) total albedo; (6) thermal radiation radiometry in the 8–14 micron region; (7) CO_2 band intensities at 0.8, 1.05, 1.20, 1.43, and 1.60 microns; (8) continuous spectral-energy distribution in the 0.75–1.6 micron region; (9) absence or weakness of CO_2 bands in the 9–11 micron region; (10) H_2O abundance from the 1.13-micron band; (11) microwave radiation 3–10 cm; (12) evidence of slow rotation; (13) occultation of Regulus; (14) photometric interpretation of micrometer diameter.

The same cannot be said of other published models, based on partial evidence, a priori reasoning, and intuition, and sometimes even on arbitrary adaptation of contradictory observational data.

The various arguments on which our model is based are described in the review article by Öpik [1961]. Here some details not given in the review will be mentioned.

6. *Estimates of pressure.* The pressure at an effective reflecting level of the Venus clouds can be estimated from the strength of the CO_2 bands. In the 'normal' or square-root region of band absorption, the equivalent width of the band, resulting from the combined absorption in all its rotational lines, is given approximately by equation 6, with

$$\Delta\lambda = \int A \, d\lambda \quad (17)$$

Here $A = A(\lambda)$ is the fractional absorption, or depression in spectral intensity, of the original continuous background, and the integral is taken over the entire band width within certain effective limits of wavelength from λ_1 to λ_2 . For strong absorption when the lines begin to over-

lap, equation 6 fails, but then its substitute can be used,

$$\Delta\lambda = (\lambda_2 - \lambda_1) \cdot \{1 - \exp [-C(pm)^{1/2}/(\lambda_2 - \lambda_1)]\} \quad (18)$$

In both cases $\Delta\lambda$ is a pure function of the product pm , pressure broadening being achieved equally by the variation of either pressure or of mass; hence, generally,

$$\Delta\lambda = f(pm) \quad (19)$$

If the function $f(pm)$ is empirically determined in the laboratory, the product pm can be inferred from an observed value of $\Delta\lambda$ without explicitly using equation 6 or 18. Because of the difficulty of restoring the original continuous background, there is some uncertainty in estimating A and $\Delta\lambda$ in equation 17, so that subjective differences may be considerable.

For very weak lines, when self-absorption is negligible, instead of equations 6, 18, and 19, the linear relation

$$\Delta\lambda = m \times \text{constant} \quad (20)$$

holds; the equivalent width is then proportional to the absorbing mass column per square centimeter and independent of pressure. All these equations are valid only in the absence of re-emission, which certainly is true for Venus for $\lambda < 2$ microns.

In a planetary atmosphere, the absorbing mass at reflection from a level of constant pressure, for incidence of the sun rays at an angle i and reflection at an angle z with the vertical, is

$$m = (\sec i + \sec z)m_0 \quad (21)$$

where m_0 is the mass of the absorbing constituent (CO_2) per square centimeter in a vertical column. For summary reflection from a planetary disk, usually $(\sec z)_{\text{av}} = 2$ is assumed and

$$\bar{m} = 4m_0 \quad (22)$$

However, there is no good definition of the mass to be used in the interpretation of spectroscopic observations when they refer to integrated conditions or the entire disk of a planet. If $\sec i + \sec z$ is weighted according to projected area, $\cos z \, dz$, and illumination, which can either be set proportional to $\cos i$, or be assumed to be constant, different averages are obtained as shown in Table 1.

Most of the published spectral observations of Venus, including all those discussed in this paper, have been made near quarter or slightly gibbous phase. Despite the diversity of the averages, it appears that $M_{\text{av}} = 4$ as adopted in equation 22 is an acceptable median value for a thin atmosphere, especially in equation 19, where the square of the average square of M is to be used; this is slightly less than the plain average.

For an optically thick atmosphere, where the effective reflecting level at oblique observation is higher and its relative air mass smaller than $\sec i + \sec z$, M_{av} must be less than 4. For Venus it appears safe to assume $M_{\text{av}} = 3$,

$$\bar{m} = 3m_0$$

Further, from the nature of pressure broadening it follows that, in a planetary atmosphere of constant composition, one-half of the bolometric pressure must be used in equations 6, 18, and 19.

Hence, in the region of nonlinear absorption and for integrated light from a planetary atmosphere, if p_0m_0 is the laboratory product that yields the same value of $\Delta\lambda$ as that of the planet, with regard to the exact form of the function in equation 19, the total pressure of the atmosphere p_0 and the mass load m_0 of the absorbing constituent (CO_2) at the reflecting level are determined through the product

$$p_0m_0 = \frac{2}{3}p_1m_1$$

TABLE 1. Averages of $M = \sec i + \sec z$

Portion of Disk Illumination	Full Phase				Quarter Phase		
	All		Central Slit		All		Equator
	Constant	$\cos i$	Constant	$\cos i$	Constant	$\cos i$	Constant
M_{av}	4.00	3.00	3.14	2.80	4.00	4.71	3.86
$[(\sqrt{M})_{\text{av}}]^2$	3.56	2.88

eliminating m_0 with the aid of the hydrostatic equation

$$p_0 = g m_0 / X$$

where g = acceleration of gravity and X = fraction by weight of CO_2 in the atmosphere, using an additional factor in $\Delta\lambda$ to allow for the greater pressure broadening of CO_2 by CO_2 as compared with that of CO_2 by N_2 (air) [Lambert, 1961], an expression for the atmospheric pressure, p_0 , at the planet's reflecting surface is obtained:

$$(1 + 0.3X) X p_0^2 = 0.87 p_1 m_1 g \tag{25}$$

For Venus, with $g = 833 \text{ cm/sec}^2$ as an average value, and with p_0 , p_1 in atmospheres, m_1 in grams per square centimeter, equation 25 becomes

$$1390(1 + 0.3X) X p_0^2 = p_1 m_1 \tag{26}$$

In the square-root region of band absorption, equation 26 is equivalent to that derived by Lambert [1961] except for the use of equation 25 instead of equation 22.

In the region of linear absorption $m_0 = \frac{1}{3} m_1$, according to equation 23, and for Venus

$$3650 X p_0 = m_1 \tag{26a}$$

For the near infrared 0.78–0.87 micron bands, Lambert's [1952] estimate sets $m_1 = 100 \text{ g/cm}^2$, $p_1 = 10 \text{ atm}$, and Herzberg's [1952] corre-

sponds to $m_1 = 400 \text{ g/cm}^2$ at $p_1 = 1 \text{ atm}$. Hence, from equation 26,

$$p_0 = 0.85[X(1 + 0.3X)]^{-1/2} \quad (\text{Dunham})$$

and

$$p_0 = 0.54[X(1 + 0.3X)]^{-1/2} \quad (\text{Herzberg})$$

in atmospheres. For this kind of estimate, the discrepancy is by no means large.

From the tracings in the laboratory and of Venus published by Kuiper [1952] and Herzberg [1952], the author has made independent estimates of $\Delta\lambda$ for the CO_2 bands at 1.05, 1.20, 1.43, and 1.60 microns as listed in the Appendix and has calculated the pressures. Also, from photometric data in the ultraviolet (0.365 micron), the pressure of the Rayleigh-scattering atmosphere overlying the ultraviolet absorbing layer (albedo 0.15 only) was estimated [Öpik, 1956] as given by equation 27, which allows for the specific scattering power of CO_2 as compared with the rest, assumed to be N_2 [Van de Hulst, 1952]

$$p_0 = 0.106 - 0.032X \tag{27}$$

atmospheres.

Table 2 contains a summary of the various determinations of the pressure. The sources for the estimate of $\Delta\lambda$ are Öpik (see Appendix), Dunham, and Herzberg. The bands at 1.05 microns are weak, and the pressures may be in-

TABLE 2. Pressure at the Continuous Background Reflection Levels of Venus in Different Wavelengths

λ , microns Method	0.365 Albedo	0.80* CO_2	1.05* CO_2	1.05† CO_2	1.20* CO_2	1.43* CO_2	1.60* CO_2
X ‡	p_0 , atmospheres						
0.8	0.080	...	0.9	0.34	0.30	0.04	0.09
0.4	0.092	...	1.2	0.68	0.45	0.06	0.13
0.2	0.100	...	1.8	1.36	0.66	0.09	0.19
0.8	...	0.85
0.4	...	1.27
0.2	...	1.84
0.8	...	0.54	0.44	0.12
0.4	...	0.81	0.60	0.17
0.2	...	1.17	0.96	0.25

* Nonlinear (square-root) interpolation or extrapolation used.
† Linear extrapolation.
‡ Fraction of CO_2 by weight.
D., Öpik; D., Dunham; H., Herzberg.

intermediate between those obtained with the linear and the nonlinear formulas.

For the 1.60-micron bands, *Kaplan's* [1961] relation yields 0.15, 0.22, and 0.30 atmosphere, respectively, for Herzberg's estimate of $\Delta\lambda$ in the Venus spectrum calibrated by another source. The agreement with the three bottom values in the last column of the table is satisfactory.

7. *The two reflecting levels.* The calculated pressures appear to fall into two distinct groups: the high values for the 0.8–1.20 micron interval; and the low values at 1.43, 1.60, and 0.365 micron. At 1.20 microns the pressure is intermediate but nearer to the high value. The low value for the ultraviolet is supported by polarimetric observations [*Gehrels*, 1960]; if interpreted as due to Rayleigh scattering above a cloud layer, these indicate a pressure of the order of 0.04 atm only [*Öpik*, 1961]. On the other hand, photometry by *Ross* [1927] of the Venus dark bands indicates that the optically determined mass of the Rayleigh-scattering atmospheric 'valley' between the dark bands is 2.3 times greater in the blue, at 0.44 micron, than in the ultraviolet at 0.365 micron [*Öpik*, 1956, p. 44]. The total pressures above the reflecting levels may be in the same ratio, indicating a gradual transition and increase of pressures and air mass from the ultraviolet toward the longer wavelengths.

All this consistently points to the presence of two main reflecting and radiating layers on Venus: an upper layer of haze, opaque in the ultraviolet and in the infrared beyond 1.4 microns, but transparent between 0.45 and 1.2 microns, to which apparently the layer that emits thermal radiation [*Sinton and Strong*, 1960] also belongs; and a lower layer of the visible clouds, from which the visible and near infrared light between 0.45 and 1.2 microns is reflected. The upper haze may consist of particles in the diameter range 0.4 to 0.6 micron, absorbing in the ultraviolet and thermal infrared, purely scattering in the other spectral regions. By applying the theory of scattering and absorption [*Van de Hulst*, 1946, 1949, and 1952], it can be shown that particles of this size, of a geometrical coverage $\tau_0 = 5 \text{ cm}^2/\text{cm}^2$ (equation 14), and of a complex refractive index $1.25 - ik$, will form an opaque and dark veil at 0.365 micron and in the thermal infrared (8–14 microns) if k is not negligible (0.10–0.25), and a

transparent layer in the visual and near infrared when $k = 0$ (no absorption) because of forward scattering at large values of the scattering parameter

$$\alpha = 2\pi r/\lambda$$

Forward scattering has little effect on the intensity of the illuminating beam, so that lower clouds are illuminated and visible if the upper haze were transparent.

In the 1.4–1.6 micron region, the low pressure (Table 2) indicates that the reflecting level must be that of the upper haze. With $2r = 0.4$ micron, from equation 28, $\alpha \cong 1.6$ at $\lambda = 1.4$ micron and $\alpha \cong 1.2$ at $\lambda = 1.6$ microns; the first corresponds to pronounced forward scattering and a transparent veil; the second case to omnidirectional scattering and nearly opaque veil with high albedo [cf. Table 10, *KK*, 1952, and Table 5, *Van de Hulst*, 1952].

Another possibility could be that of particles of a larger diameter, scattering in the forward direction in all wavelengths shorter than 1.6 microns (effective omnidirectional cross section $0.25\pi r^2$ as from refraction and reflection); however, the reflecting layer at 1.4–1.6 microns could be raised to the level of the upper haze only at the expense of true absorption and corresponding low albedo, which is not the case. From a restoration of the continuous spectrum of ground in the Venus spectrum in published tracings [*Kuiper*, 1947, 1952], the authors find, indeed, some indication of a slight decrease of the reflecting power in the region of 1.6 microns.

Wavelength, microns	0.75	0.8	1.0	1.3	1.5
Relative albedo	1.00	0.95	0.92	0.82	0.78

This alone, however, is insufficient to account for the higher level at 1.6 microns and, possibly, is not in too much contradiction with *KK* [1947] statement that 'outside the CO_2 bands the reflection of Venus is nonselective (white) for the interval $\frac{3}{4} - 2\mu$.'

The temperature of 285°K , determined from the rotational lines of CO_2 in the 0.8-micron region [*Chamberlain and Kuiper*, 1956], refers to the same bands from which abundance and pressure were derived (Table 2, third column). The temperature refers to the center of mass of the absorbing column and may thus be the

temperature at two-thirds of a scale height, or about 5 km above the lower cloud level.

The radiometric temperature of 234°K [Sinton and Strong, 1960] in the 8-14 micron region must belong to a higher level; its identification with the upper haze layer yields a plausible temperature gradient for the indicated pressure ratio. The layer is opaque to thermal radiation, which can only be due to true absorption ($\alpha \cong \kappa$), cross section for scattering $\sim 10^{-4}$ of geometrical cross section, equations 28 and 16). A former proposal by Öpik [1956], to consider the observed radiation as gray emission of emissivity 0.25 from the warm cloud surface, cannot be upheld, as in this case the CO₂ bands at 10.4, and 12.6 microns would have appeared in emission (as verified by calculations, allowing for the Boltzmann factor of these temperature-dependent bands). There is almost no trace of these bands in the radiometric records [Sinton and Strong, 1960, Figs. 13, 15], which agrees with the hypothesis of an upper haze layer, especially if the temperature distribution immediately above it is isothermal as in a stratosphere. Also, the amount of CO₂ above the haze is small, and the excitation of the lower levels of these bands is weak, owing to the low temperature.

Vertical cross section of the atmosphere of Venus. The recent observations of the occultation of Regulus by Venus [Menzel and de Vauvergne, 1960] provide a well-determined point in the Venus upper atmosphere, with the temperature-to-molecular-weight ratio and its variation with altitude being observed; this variation must be due to temperature alone, and not to molecular weight. The pressure at the occultation level, 2.6 dynes/cm², corresponds to an altitude of 90 km on earth, where the mean molecular weight is still essentially constant; the same can be expected to hold for the occultation level of Venus.

The altitude of the occultation level was estimated to be 65 ± 7 km above the 'micrometer limb.' However, the limb cannot be simply identified with a cloud or haze top. From the theory of contrast, and the distribution of surface brightness in a scattering atmosphere at the edge of a planet's disk [Öpik, 1961], the effective limb of Venus is estimated to correspond to the level where the vertical transmission coefficient of the atmosphere equals $1 - 0.0036$.

The micrometer limb is thus set at 1.34 scale heights above the haze level.

From the ratio of pressures, p_2/p_1 , and the difference in altitude, $h_2 - h_1$, between the occultation and upper haze levels, the average scale height

$$H = kT/\mu g \quad (29)$$

is derived, as defined by

$$p_1/p_2 = \exp [(h_2 - h_1)/H] \quad (30)$$

Here μ = molecular mass, k = Boltzmann constant. Equation 30 yields the harmonic mean value of the scale height and, thus, of the temperature, which, with sufficient precision, can be assumed to equal the mean temperature, \bar{T} , over the range $h_2 - h_1$.

By applying the relations of hydrostatic equilibrium either as in equation 30 or in their adiabatic form (in the aeolosphere and the troposphere), self-consistent models of the Venus atmosphere have been constructed from the available points of reference as shown in Table 3 and Figure 1. The models literally in-

TABLE 3. Models of Venus Atmosphere

S = surface; C = visible cloud level; T = rotational temperature level; R = thermal radiation and upper ultraviolet haze level; L = micrometer limb; B = occultation level; X = fraction of CO₂ by weight; h = altitude, in kilometers, above C; p = pressure, atmospheres; T = temperature, °K; dT/dh = vertical lapse rate, deg/km.

Level	h	T , °K	p	dT/dh
I. $X = 0.80$; $\bar{T}(21-86 \text{ km}) = 288^\circ\text{K}$				
S	-22	570	4.3	-10.5
C	0.0	340	0.58	-10.5
T	+5.2	285	0.29	-10.5
R	13.3	234	0.080	...
L	20.9
B	86	280	2.5×10^{-6}	+2.8
II. $X = 0.40$; $\bar{T}(27-92 \text{ km}) = 245^\circ\text{K}$				
S	-25	570	5.4	-9.2
C	0.0	344	0.82	-9.2
T	+6.4	285	0.41	-9.2
R	17.6	234	0.092	...
L	26.6
B	92	236	2.5×10^{-6}	+2.4
III. $X = 0.20$; $\bar{T}(31-96 \text{ km}) = 223^\circ\text{K}$				
S	-26	570	7.0	-8.7
C	0.0	346	1.16	-8.7
T	+7.0	285	0.58	-8.7
R	21.4	234	0.100	...
L	31.3
B	96	214	2.5×10^{-6}	+2.1

corporate all the observational data, without modification and without allowing for the possible errors. The pressure at level C is a certain estimated average of the values of Table 2 for the 0.8–1.2 micron range; that at level R is assumed equal to the photometric value at 0.365 micron.

In Figure 1 smooth curves are drawn through the reference points of Table 3; the bulge at $h = 50$ km is drawn so as to satisfy the condition of average scale height and average temperature over the appropriate range.

As to choice, model I with 80 per cent CO₂ gives the best agreement in pressure at the 0.365 and 1.4–1.6 micron levels (Table 2) and appears to claim some preference on this, rather weak, ground. Kaplan [1961] gives preference to a low CO₂ content on the ground that its low upper atmospheric temperatures seem to be in better accord with radiative equilibrium at an effective radiative temperature of 234°K as for the observed thermal radiation. However, solar ultraviolet radiation must penetrate to the occultation level and contribute to its heating, by way of dissociation of CO₂ and its subsequent recombination in three-body collisions; the equilibrium temperature may then be expected to be well above 234°K (as indicated by a sample calculation), and a temperature around 280°K as in model I ($X = 0.80$) appears to be quite possible. The matter requires further study.

9. *Note on water-vapor abundance.* The November 1959 balloon flight yielded $m_1 = 5.3 \times 10^{-3}$ g/cm² of H₂O equivalent at $p_1 = 1$ atm (Strong, private communication, 1960) for Venus (phase angle 80°, slightly gibbous), in our notation. 'If the pressure is less on Venus (as we expect) than it was on the occasion of calibration at the stratobowl in South Dakota, this modulation would correspond to a higher water content. The increase would be small unless the pressure is below 0.1 atm since the modulations were nearly in the linear region rather than the square-root-law region. The water vapor in the

part of the optical path in the earth's atmosphere above the gondola was negligible.' (Strong, private communication, 1960.)

If this is so, equations 23 and 20 apply with $m_0 = 1.8 \times 10^{-3}$ g/cm², close to the value derived by Strong.

From the set of pressures of Table 2 it appears that the 1.13-micron band of H₂O originates at level C, or possibly slightly above near level T (temperature 285°K). Assume the pressures for this level as given in Table 2 and equation 24 as for the square-root region; the curve of growth, the water-vapor amount, turn out to be as in Table 4.

TABLE 4. H₂O Amount (m_0) and Densities (δ) from the Moore-Ross Balloon Flight

X (CO ₂)	0.8	0.4	0.2
p_0 , atm	0.29	0.41	0.58
m_0 , g/cm ²	12×10^{-3}	8.5×10^{-3}	6.0×10^{-3}
δ , g/cm ³	1.5×10^{-8}	8.9×10^{-9}	5.7×10^{-9}

These figures represent improbable limits, although they are smaller than the limits proposed by Sagan [1961] in his very stimulating and imaginative article. With the very temperature distribution as in Figure 1, however, the moisture content of a mixed atmosphere at all levels will remain far below the saturation value (actually 1 to 3 per cent saturation). It is hard to see how ice-crystal (cirrus) clouds could form under such conditions.

Moreover, observations are against the existence of such clouds; in the ultraviolet near the edge of the disk near the limb is even nearly as black as would be expected from ice-crystal clouds observed obliquely, every detail corresponding to a low albedo of 0.15 to 0.20 [cf. Ross, Öpik, 1956].

APPENDIX

Estimates of equivalent widths of CO₂ bands in Kuiper's [1952] and Herzberg's [1952] spectra are shown below.

Wavelength, microns	1.038, 1.050, 1.065	1.206, 1.220	1.43	1.54, 1.60, 1.66
	4.2×10^{-3}	Venus $\Delta\lambda$, microns 7.3×10^{-3}	1.20×10^{-2}	3.00×10^{-2}
		Laboratory $\Delta\lambda$, microns		
$p = 0.12$ atm, $m = 33$ g/cm ²	Nil	1.5×10^{-3}	1.33×10^{-2}	2.00×10^{-2}
2.0	35	3.5×10^{-3}	2.70×10^{-2}	5.4×10^{-2}
1.0	272	1.58×10^{-2}	2.35×10^{-2}	8.4×10^{-2}
1.0	17	1.5×10^{-3}	1.33×10^{-2}	3.8×10^{-2}

acknowledgment. The research for this paper supported by grant NsG-58-60 from the National Aeronautics and Space Administration.

REFERENCES

- Pett, A. H., *Astrophys. J.*, **133**, 281-293, 1960.
 Amberlain, J. W., and G. P. Kuiper, *Astrophys. J.*, **124**, 399-405, 1956.
 Ham, T., Spectroscopic observations of the planets at Mt. Wilson, in *The Atmospheres of the Earth and Planets*, edited by G. P. Kuiper, University of Chicago Press, Chicago, 434 pp., 1952.
 Lythe, W. E., *Smithsonian Physical Tables*, 6th ed., Smithsonian Institution, Washington, D. C., 1956.
 Fels, Th., *Lowell Observatory Bull.*, **4**, 300, 1960.
 Don, J. E., and R. J. McEwan, Observations of Venus at 8.6 mm wavelength, in *I. A. U. Paris Symposium on Radio Astronomy*, edited by N. Bracewell, Stanford University Press, Stanford, 612 pp., 1959.
 Berg, G., Laboratory absorption spectra obtained with long paths, in *The Atmospheres of the Earth and Planets*, edited by G. P. Kuiper, University of Chicago Press, Chicago, 434 pp., 1952.
 an, L. W., The structure and CO₂ content of the Venus atmosphere, Rand Corporation report, prepublication copy, January 1961.
 er, G. P., *Astrophys. J.*, **106**, 251-254, 1947; *Donald Observatory Contrib.* **140**, 1947.
 er, G. P., Planetary atmospheres and their origin, in *The Atmospheres of the Earth and Planets*, edited by G. P. Kuiper, University of Chicago Press, Chicago, 434 pp., 1952.
 Bolt-Börnstein, *Physical Tables*, 6th ed., vol. part 2, Springer-Verlag, Berlin, 1951.
 R. J., *Smithsonian Meteorological Tables*, 6th ed., Smithsonian Institution, Washington, D. C., 1951.
 Mayer, C. H., *Astron. J.*, **64**, 43-45, 1959.
 Mayer, C. H., T. P. McCullough, and R. M. Sloanaker, *Astrophys. J.*, **127**, 1-10, 1958; *Proc. IRE*, **46**, 260-266, 1958.
 Mayer, C. H., T. P. McCullough, and R. M. Sloanaker, *Astron. J.*, **65**, 349, 1960.
 Menzel, D. H., and G. de Vaucouleurs, *Astron. J.*, **65**, 351, 1960.
 Öpik, E. J., *Irish Astron. J.*, **1**, 37-47, 1950; *Armagh Observ. Leaflet* **4**, 1950.
 Öpik, E. J., *Geophys. Bull. Dublin Inst. Adv. Studies*, **8**, 1-14, 1953.
 Öpik, E. J., *Irish Astron. J.*, **4**, 37-48, 1956; *Armagh Observ. Leaflet* **43**, 1956.
 Öpik, E. J., The atmosphere and haze of Mars, *J. Geophys. Research*, **65**, 3057-3063, 1960.
 Öpik, E. J. Atmosphere and surface properties of Mars and Venus (preliminary abstract, *Am. J. Phys.*, **28**, 618-622, 1960), in *Progress in the Astronautical Sciences*, edited by S. F. Singer, vol. 1, North Holland Publishing Company, Amsterdam, 1961.
 Ross, F. E., *Astrophys. J.*, **68**, 57-92, 1928; *Mt. Wilson Observatory Contrib.* **363**, 1928.
 Sagan, C., *Science*, **133**, 849-858, 1961.
 Sagan, C., K. M. Siegel, and D. E. Jones, *Astron. J.*, **66**, 52-53, 1961.
 Sinton, W. M., and J. Strong, *Astrophys. J.*, **131**, 470-490, 1960.
 Strong, J., M. D. Ross, and C. B. Moore, *J. Geophys. Research* (abstract), **65**, 2526, 1960.
 Van de Hulst, H. C., *Recherches astron. de l'observatoire d'Utrecht*, **11**, part 1, 1-87, 1946; part 2, 1-50, 1949.
 Van de Hulst, H. C., Scattering in the atmospheres of the earth and planets, in *The Atmospheres of the Earth and Planets*, edited by G. P. Kuiper, University of Chicago Press, Chicago, 434 pp., 1952.

(Manuscript received May 19, 1961.)



JOINT SCIENTIFIC ADVISORY GROUP TO THE METEOROLOGICAL ROCKET NETWORK¹

2821

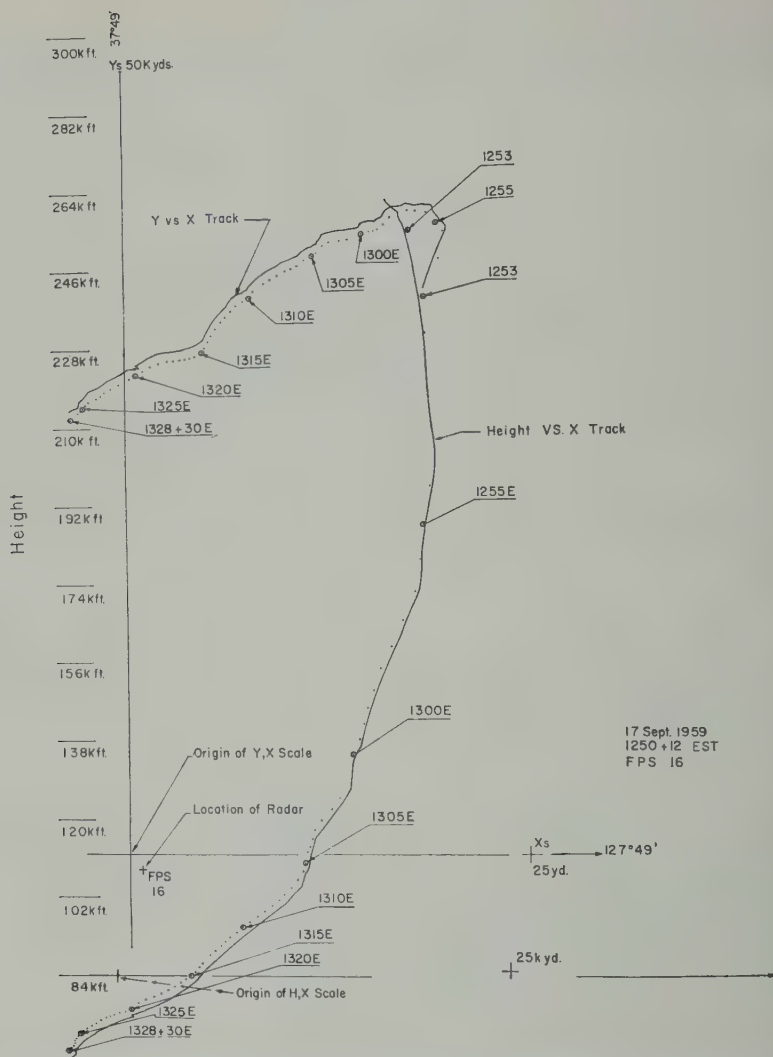


Fig. 2. Reproduction of a radar plot used to compute winds. An extremely good track is illustrated.

provide a broad picture of the usefulness and potential of a rocket network as a new meteorological tool.

The Meteorological Rocket Network commenced operations in the fall of 1959 with two stations: Point Mugu, California, and Fort Churchill, Canada. During the first year five more stations were added: Wallops Island, Virginia; Cape Canaveral, Florida; Tonopah, Nevada; Fort Greely, Alaska; and White Sands, New Mexico. Two additional stations were added in the fall of 1960 at Eglin AFB, Florida, and Barking Sands, Kauai, Hawaii (Fig. 1). Future expansion will probably take place at a

slower rate because of the lack of suitable launching areas in the western hemisphere.

The basic observational schedule of the network is the firing of one rocket per day at each station during four midseason months. In addition to these are firings designed to collect special research data. In practice this schedule will be modified as a result of the network's dependence upon range support. The actual schedule consisted of firings on each weekday (if suitable rockets were available) during 4-week periods beginning about the 15th of January, April, July, and October.

Two basic rocket motors are used to

TABLE 1. Time (ΔT) and Height (ΔH) Required for .012-Inch Nylon Chaff to Attain a Velocity of 5 Per Cent of the Wind Velocity (case A) and 5 Per Cent of Its Initial Velocity (case B) when Dropped from 90 Kilometers and Entering a Layer of Wind Velocity W at Height H_w (The velocity of the chaff upon entering this layer is V .)

Case A		Case B	
$W = 130$ m/sec	$V = -160$ m/sec	$W = 0$	$V \neq 0$
ΔH , km	ΔT , sec	ΔH , km	ΔT , sec
16.95	87.5	16.10	80.2
12.40	57.2	11.65	50.0
7.80	44.0	7.15	37.0
3.55	32.0	3.15	26.8
1.00	21.7	0.85	18.0
0.30	11.1	0.20	7.4

instrument package into the upper atmosphere. *Jenkins and Webb* [1959] have discussed characteristics of the Loki, and *Webster, Roberts, and Donnell* [1960] have written a complete report on the Arcas.

Two types of wind sensors were used during the first year of network operations—radar active parachutes and chaff. A 15-ft parachute was used with the Arcas rocket, and, particularly at Tonopah, Nevada, an 8-ft parachute was used with the Loki I. Chaff, with a few exceptions, was used only with the more powerful Loki II. Parachutes have the advantage of presenting discrete targets which can be tracked as long as they remain within the line of sight of the radar and are capable of carrying an instrument package. However, they fall at a faster rate than most types of chaff and consequently are more sensitive to fluctuations of the wind. Chaff has the disadvantage of diffusing as it falls, and usually cannot be tracked as a coherent target up to the levels of balloon winds as can the parachute.

Temperature sensing and telemetering packages for use with the Arcas rocket are being developed [Clark, 1961]. The Gamma package, in which the sensor is a bead thermistor and the transmitter is the ground equipment, has been used extensively thus far. A similar package for the DMQ-6 is under development for use with the GMD-2 equipment. Further discussion of the characteristics of the wind and tem-

perature sensing devices appears in another section.

DATA REDUCTION AND EVALUATION

Winds. The method of computing winds from the trajectory of rocket-released sensors is essentially the same as that used for a standard Pibal observation except that radar tracking replaces optical tracking. There are many ways by which the radar positional data can be converted into winds, but in this report the winds were determined by a graphical technique. A plotting board (Fig. 2) continuously and automatically records, in cartesian coordinates, the position of the target in a horizontal plane and a vertical plane. The wind is computed from the horizontal displacement during each minute and is assumed to have occurred midway between the heights of the target at the beginning and end of the minute.

Some important factors which do not apply to balloon winds must be taken into account in reducing and evaluating rocket winds. One factor is that the target, upon release into the atmosphere, has a velocity close to the ballistic velocity of the rocket. Until the target approaches equilibrium with the wind, a correction must be applied to convert target velocity into wind velocity. This correction factor will soon be determined for the 15-ft parachute. To circumvent the possibility of reporting target velocities as winds during this transition period, the authors made subjective determinations of the

TABLE 2. Estimated Error in Wind Speed and Height Based upon the Departure from the Mean Value as Determined by FPS-16 and Mod II Radars

Date and Location	Error		
	E-W Wind, m/sec	N-S Wind, m/sec	Height, km
Canaveral			
8/3/60	0.65	0.68	0.136
8/11/60	0.67	0.64	0.065
Wallops Is.			
7/20/60	0.98	0.99	0.686
8/3/60	2.03	1.66	0.249
8/4/60	3.09	1.87	0.095
8/9/60	2.22	1.28	0.240

TABLE 3. Approximate Fall Velocities in Meters per Second of Various Radar Wind Targets at Selected Altitudes

Target	60 km	45 km	30 km
15-ft parachute with payload	140	45	13
8-ft parachute		40	12
.010-in. copper chaff	75	43	23
.012-in. nylon chaff	20	10	3

levels at which the targets approached equilibrium with the wind. Although this procedure may have resulted in the loss of some good data, it was considered more desirable than the alternative of presenting misleading wind values. Furthermore, in order to apply the corrections it is necessary to have data at a much smaller time increment than 1 minute, and resolution of the trajectory to the necessary degree from the radar plots is not always practical.

Most chaff is not significantly affected by the wind equilibrium problem at the altitudes realized during the first year [Barr, 1960]. Table 1 lists the height increment (ΔH) and the time increment (ΔT) required for .012-inch nylon chaff to approach equilibrium with the wind under two sets of extreme conditions. The conditions in case A were that the chaff reach a velocity 95 per cent of the wind velocity after entering a layer with a wind velocity W of 130 m/sec at an initial velocity V of 160 m/sec in a direction *opposite* to the wind. In case B, the conditions were that the chaff attain 5 per cent of its initial velocity after entering a layer of no wind. The absolute value of the initial velocity is not significant as long as it is not equal to zero. In both cases the chaff was assumed to have been dropped from 90 km. At height H_w it enters the specified wind field at the specified initial velocity.

The implication of this analysis is that .012-inch nylon chaff approaches the wind velocity very rapidly even under extreme conditions. At 75 km it takes approximately $\frac{1}{2}$ minute to attain 95 per cent of the wind velocity regardless of the initial difference between the wind and target velocity. Since the winds are averaged over 1-minute intervals, no serious error occurs in the wind values below 75 km. Comparison of the two cases in Table 1 reveals that there is no

significant variation in the time or height it takes the chaff to approach an equilibrium velocity. This is because the horizontal acceleration is proportional to the velocity difference between the target and the wind.

The foregoing discussion involves the behavior of the target in response to the forces acting upon it. Another factor which affects the accuracy is the tracking error of the radar equipment. So many combinations of targets and radars were used during this first year that there has not yet been a complete evaluation of this aspect of data quality.

Rapp [1960] analyzed the chaff winds reported by Smith [1960] and determined that the correlation between chaff and wind velocity was about 1.5 mps while the random instrumental error of the MSQ-1A radar was about 10 mps. On six occasions in the summer of 1960, 15-ft parachute targets were tracked simultaneously by two radars, an FPS-16 and a Mod II. The FPS-16 is a very accurate missile tracking radar and the Mod II is a modification of the SCORPIO gun-laying radar.

The estimated error in the wind speed

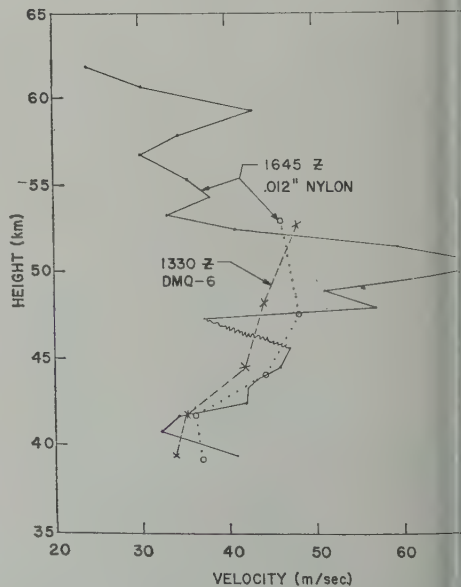


Fig. 3. Comparison of wind profiles obtained by computing 1-minute average winds from slower falling chaff (solid line) and a faster falling 15-ft parachute with a DMQ-6 attached (broken line). The dotted line is the chaff wind profile when averaged over the same height increment as the parachute in 1 minute.

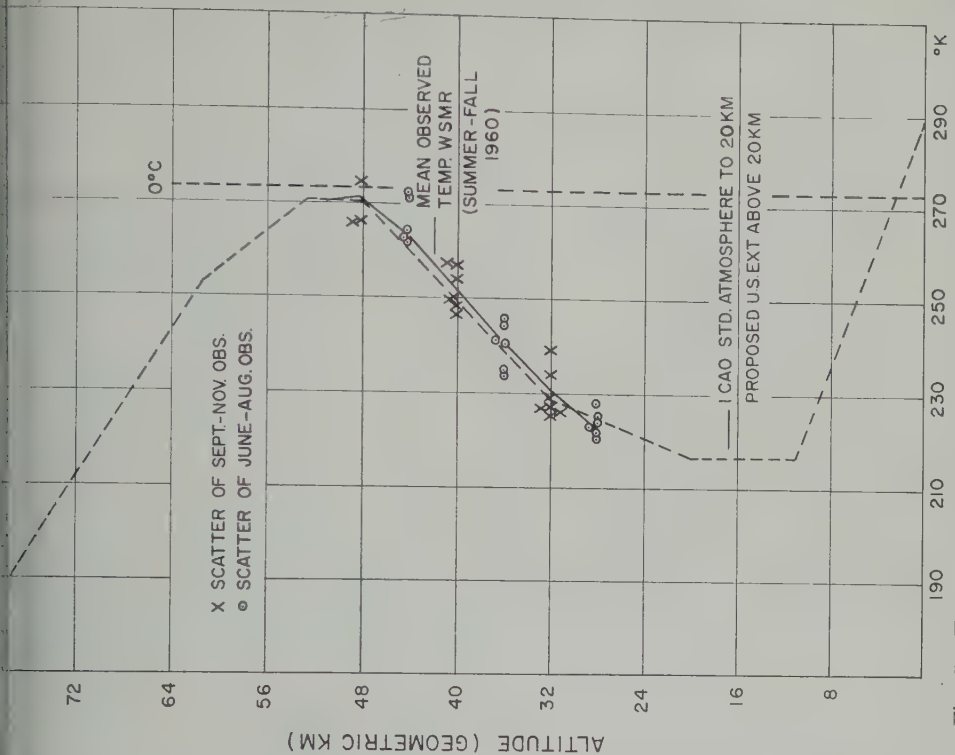


Fig. 5. Temperatures over White Sands, New Mexico, during summer and fall 1960. Plotted values have been corrected according to Table 5.

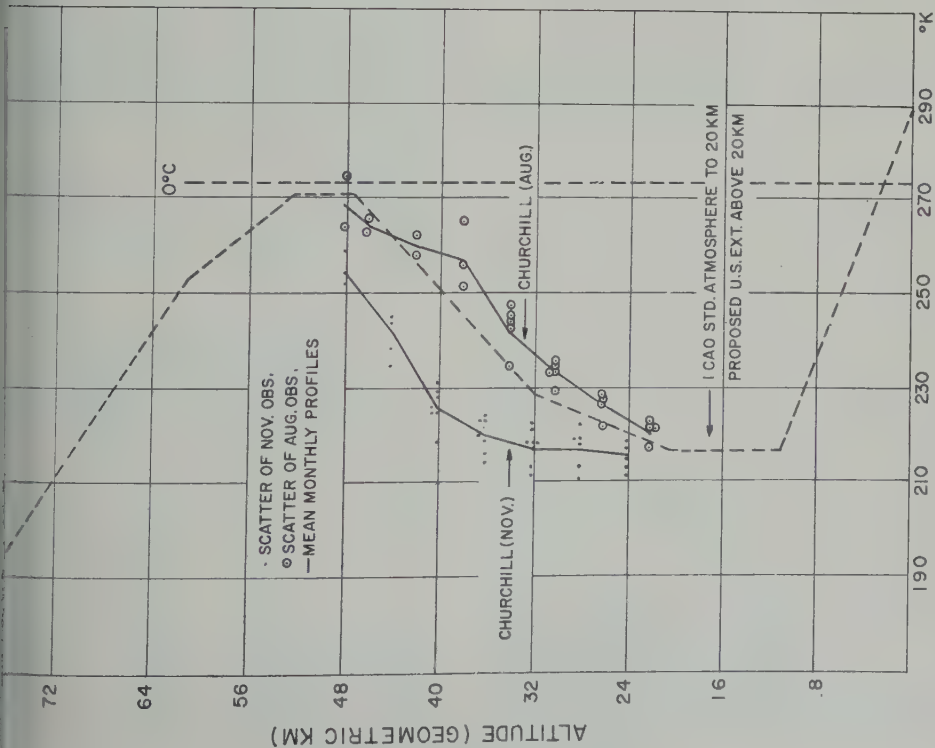


Fig. 4. Temperatures over Fort Churchill, Canada, during summer and fall 1960. Plotted values have been corrected according to Table 5.

TABLE 4. Temperature Corrections for Gamma Package Lowered by a 24-Foot-Diameter Parachute (Effect of self heating is not included.)

Alt., km	v , cm/sec	h_v , $\frac{\text{cal}}{\text{cm}^2 \text{ sec } ^\circ\text{K}}$	θ_1 , $^\circ\text{K}$	θ_2 , $^\circ\text{K}$	θ_3 , $^\circ\text{K}$	θ_4 , $^\circ\text{K}$	θ , $^\circ\text{K}$
30.0	0.95×10^3	5.095×10^{-3}	0.04	0.12	0.006	0.017	0.149
32.0	1.05×10^3	4.76×10^{-3}	0.05	0.13	0.006	0.020	0.166
34.1	1.20×10^3	4.37×10^{-3}	0.06	0.14	0.007	0.024	0.183
35.9	1.36×10^3	4.06×10^{-3}	0.08	0.15	0.007	0.028	0.209
38.0	1.58×10^3	3.66×10^{-3}	0.11	0.17	0.008	0.035	0.253
40.1	1.85×10^3	3.27×10^{-3}	0.15	0.18	0.009	0.042	0.297
41.9	2.09×10^3	2.94×10^{-3}	0.20	0.20	0.010	0.052	0.358
44.0	2.41×10^3	2.55×10^{-3}	0.26	0.24	0.012	0.065	0.447
46.1	2.75×10^3	2.18×10^{-3}	0.34	0.27	0.014	0.083	0.541
47.9	3.07×10^3	1.89×10^{-3}	0.42	0.32	0.016	0.101	0.655
50.0	3.50×10^3	1.61×10^{-3}	0.55	0.37	0.019	0.119	0.820
52.1	3.98×10^3	1.35×10^{-3}	0.71	0.45	0.022	0.142	1.040
53.9	4.45×10^3	1.16×10^{-3}	0.90	0.52	0.026	0.160	1.286
56.0	5.05×10^3	9.75×10^{-4}	1.17	0.61	0.031	0.170	1.641
58.1	5.68×10^3	8.12×10^{-4}	1.49	0.74	0.037	0.181	2.086

height computed from the mean of each pair of soundings is shown in Table 2. Sampling indicated that it was not necessary to subdivide the differences into altitude ranges. In these cases the radar error is of the order of 3 m/sec or less, compared with the 10 m/sec mentioned above for the chaff.

How well the parachute follows the wind has not been determined. Preliminary investigations indicate that, except under conditions of large shears at high altitudes, the parachute has excellent wind response characteristics.

There is one other characteristic of these soundings which must be understood in order to interpret the data properly. The winds have all been computed over a constant time interval—1 minute. However, several different types of parachutes and chaff have been used, and they possess differing fall velocities. Table 3 lists the targets and their fall velocities at selected altitudes.

Inspection of Table 3 shows that the altitude increment of the reported winds varies with both the type of sensor and with altitude. The effect this can have on the reported winds is illustrated in Figure 3. It shows wind profiles at Cape Canaveral, Florida, on August 12, 1960. The dashed line represents the wind as computed from a sounding taken with a 15-ft parachute and the solid line a sounding taken 3 hours later with .012-inch nylon chaff. At

50 km, in a zone of strong shear, the reported wind velocities differ by a factor of 2. This does not indicate either a large change in wind velocity over a short period or a lack of agreement between the two sensors, for if these winds are averaged over the same height interval in which the parachute winds represent (0.5 km line), the agreement is very good. Notice that between 40 and 45 km, where the shears are moderate, the profiles of the 1-minute averaged winds agree very well.

Temperatures. The temperature data presented in Figures 4 and 5 were obtained from Gamma instrumentation packages in rockets fired from Fort Churchill, Canada.

TABLE 5. Temperature Corrections Used for Gamma Package Lowered by a 15-Foot-Diameter Parachute (Effect of self heating is not included; symbols have same meaning as in Table 4.)

Alt., km	θ_1 , $^\circ\text{K}$	θ_2 , $^\circ\text{K}$	θ , $^\circ\text{K}$
48	1.33	0.31	1.1
46	0.95	0.27	1.0
44	0.69	0.24	0.9
42	0.50	0.21	0.8
40	0.36	0.18	0.7
38	0.26	0.16	0.6
36	0.19	0.15	0.5
34	0.14	0.14	0.4

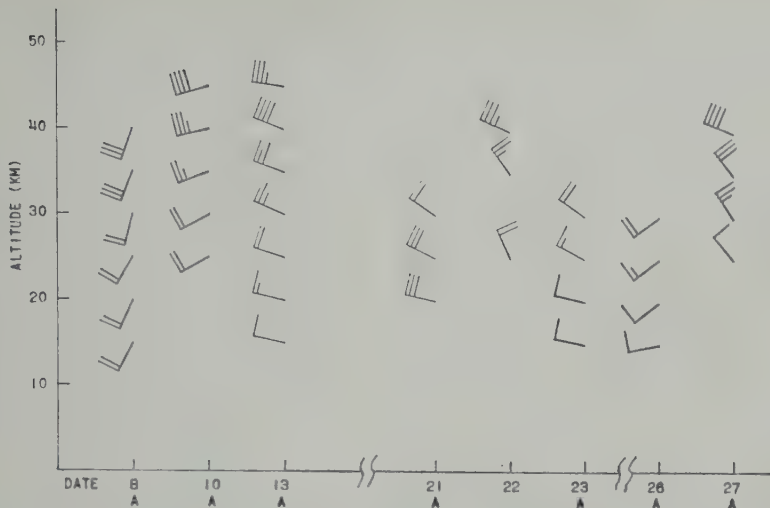


Fig. 6. Observed winds over Fort Churchill, Canada, during October 1959. A full wind barb represents 10 m/sec, and a pennant, 50 m/sec. The wind staff indicates direction with north at the top of the figure. The letter under each observation is the type of rocket and wind sensor used for the sounding according to the following code: A, Arcas rocket with 15-ft. parachute; L₁, Loki II rocket with .012 in. nylon chaff; L₂, Loki II rocket with .01-in. copper chaff; L₃, Loki II rocket with .02-in. nylon chaff; and L₄, Loki I rocket with 8-ft parachute.

the Sands, New Mexico. This package consists essentially of a standard AN/AMT-4 sonde from which the pressure sensor has been removed. The radiosonde is repackaged in a special configuration suitable for installation in the Arcas nose cone. The circuitry has been designed to minimize self heating of the tem-

perature sensor, a 10-mil ceramic bead, which is exposed directly below the package as it descends on a 15-ft metalized parachute.

Problem areas associated with the Gamma package which are being evaluated are the effects of self heating and solar radiation. Portions of the observations, particularly above 45 km when the

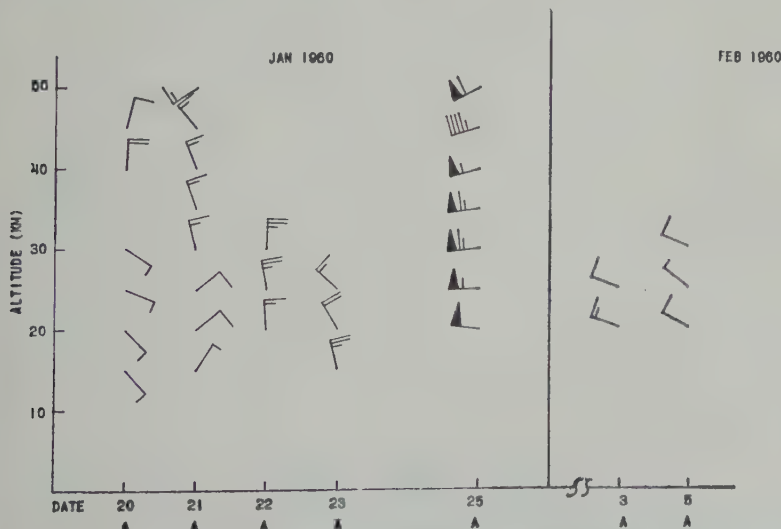


Fig. 7. Observed winds over Fort Churchill, Canada, during January and February 1960. See Figure 6 for explanation of symbols.

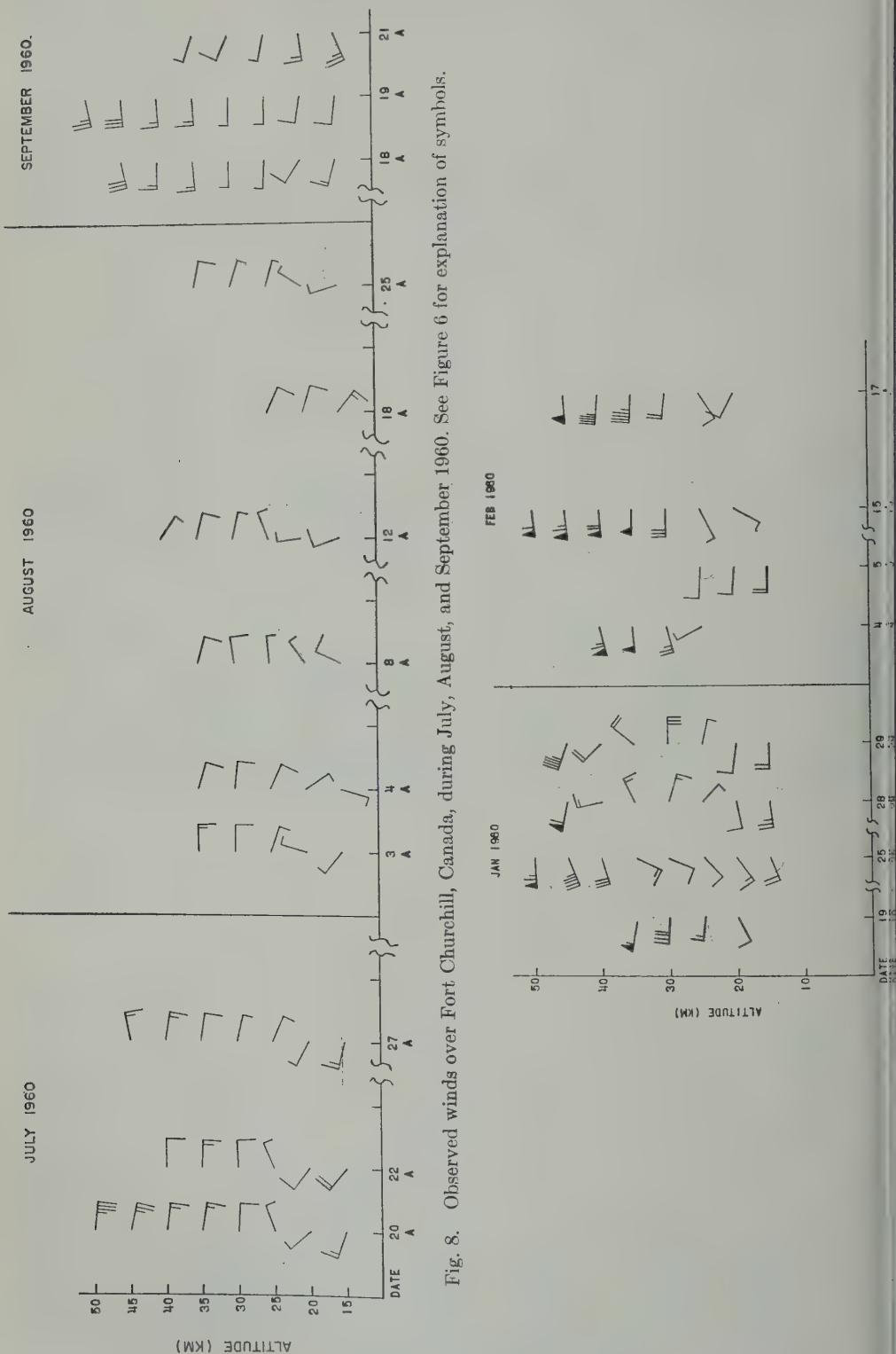


Fig. 8. Observed winds over Fort Churchill, Canada, during July, August, and September 1960. See Figure 6 for explanation of symbols.

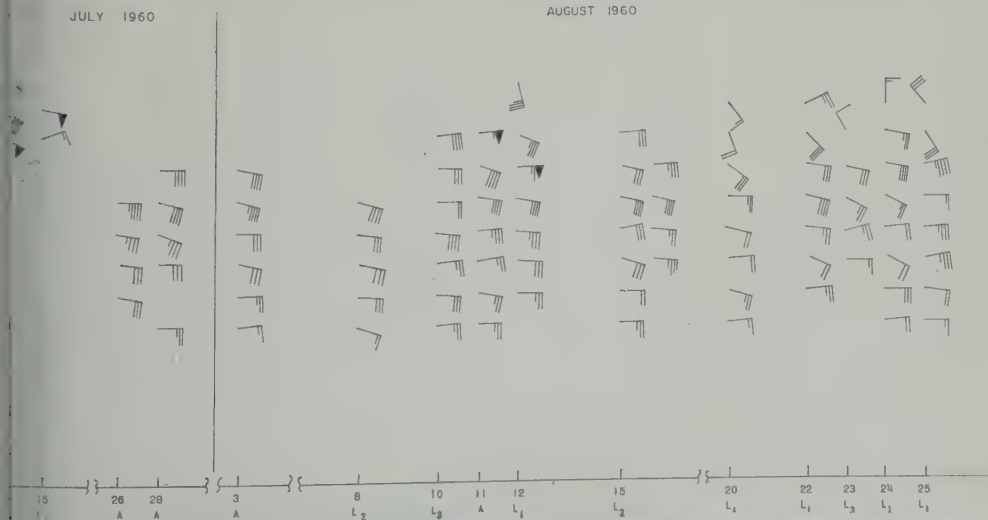
temperature continues to increase rapidly with altitude. These data are probably in error. The Research and Development Laboratory and the Missile Support Agency of the U. S. Army Signal Corps are presently working on the development and application of bead thermistors as high-altitude temperature sensors. Results of their investigation will be reported later this year [Ballard, 1961; Barr, 1961; Wagner, 1961]. In this rather limited area, there are some provisional results available at this time which are presented in Figure 4. Here v is the fall rate; k the convective heat transfer coefficient; θ_1 , θ_2 , and θ_3 the temperature increases due to aerodynamic heating, solar radiation, and long-wave radiation, respectively; θ_4 the temperature decrease due to reradiation from the element; $\theta = \theta_1 + \theta_2 + \theta_3 - \theta_4$; τ is the time constant of the element.

The fall rates used in preparing Table 4 were computed for a 24-ft parachute which has a fall rate than the 15-ft parachute currently in use. Table 5 presents the errors due to aerodynamic heating (θ_1) and solar radiation (θ_2) based on the fall rate of the 15-ft parachute. The sum of θ_1 and θ_2 was used to compute θ . θ_3 and θ_4 do not contribute significantly to the total correction at the levels treated. The values of θ in Table 5 were used to correct the temperature values plotted on Figures 4 and 5. The temperature data shown in Figure 4 are based on seven November firings and five August firings at Fort Churchill, Canada, during 1960.

Five of the November firings and two of the August firings provided usable data to altitudes above 44 km. Mean temperature profiles based on these observations are shown for November and August together with the proposed U. S. Standard Atmosphere and Extension developed during March 1961. The season variation in the mean conditions agrees very well with other types of observations made at Churchill in this altitude range.

The White Sands data shown in Figure 5 are based on seven summer (June-August) and seven fall (September-November) firings. Eight of these provide usable data above 44 km. There is no significant seasonal difference between mean summer and fall temperature profiles based on these observations. This confirms conclusions reached using other types of observational data at low-latitude stations. The mean (summer-fall) temperature profile shown in Figure 5 closely parallels that of the proposed U. S. Standard Atmosphere and Extension.

Winds in the mesosphere. The illustrations in this section will be presented with a minimum of comment, as they are self-explanatory for the most part. The figures show time cross sections of observed wind velocity in m/sec for selected stations and seasons. The winds, which are plotted at 5-km intervals, represent point winds as interpolated from a profile of linear segments connecting the computed 1-minute mean values. The target used for each sounding is identified.



10. Observed winds over Cape Canaveral, Florida, during July and August 1960. See Figure 6 for explanation of symbols.

In addition, seasonal wind profiles are presented for each station included in the network.

Some highlights of these diagrams are as follows:

(a) Figures 6, 7, and 8 show the winds at Fort Churchill, Canada, during the fall of 1959 and winter, summer and early fall of 1960. In the fall and winter (Figs. 6 and 7) there are generally increasing westerly winds with height, and the changes in wind velocity over several days can be considerable (e.g., October 8-10 and January 21-25). The wind pattern is reversed in summer, and Figure 8 shows increasing easterly winds with height. The change back to west winds occurs early in September and by mid-September the westerlies have returned to the mesosphere (30 to 80 km) over Fort Churchill.

(b) The Wallops Island, Virginia, soundings for January-February 1960 (Fig. 9) start on January 19 with a sounding which appears normal enough—westerly winds increasing with height. However, by January 25 a significant change has taken place in the lower mesosphere (30 km to 40 km). The west winds have weakened considerably and by January 28 have reversed to easterly. By February 4 the wind velocities appear to have returned to a more conventional distribution. The winds at Point Mugu during

this same period (not shown) underwent the same reversal in the lower mesosphere. Examination of radiosonde data during this period revealed that a tropical ridge moved north from south of 30°N. It was oriented E-W and the ridge line sloped northward with elevation. Indications are that at 35 km of the U. S. was south of the ridge line and reached its northern limit near the end of January.

(c) Winds at Cape Canaveral (Fig. 10) show the weakening of the high-level easterlies in late August and by the end of August westerly winds have already begun to develop above 40 km. Although they do not show up on the diagrams, considerable wind shears existed on some days. On August 25, for example, shears in the E-W wind component of .03 per sec occurred near 50 km. This value was computed from a sounding averaged more than 2 minutes in time and more than a kilometer in depth.

(d) At White Sands, New Mexico (Fig. 11 and 12), there is a very good series of soundings through the summer and into fall. In the mesosphere easterlies can be seen increasing to a maximum in July and then decreasing in mid-August. Again by mid-September the westerlies are established above 50 km.

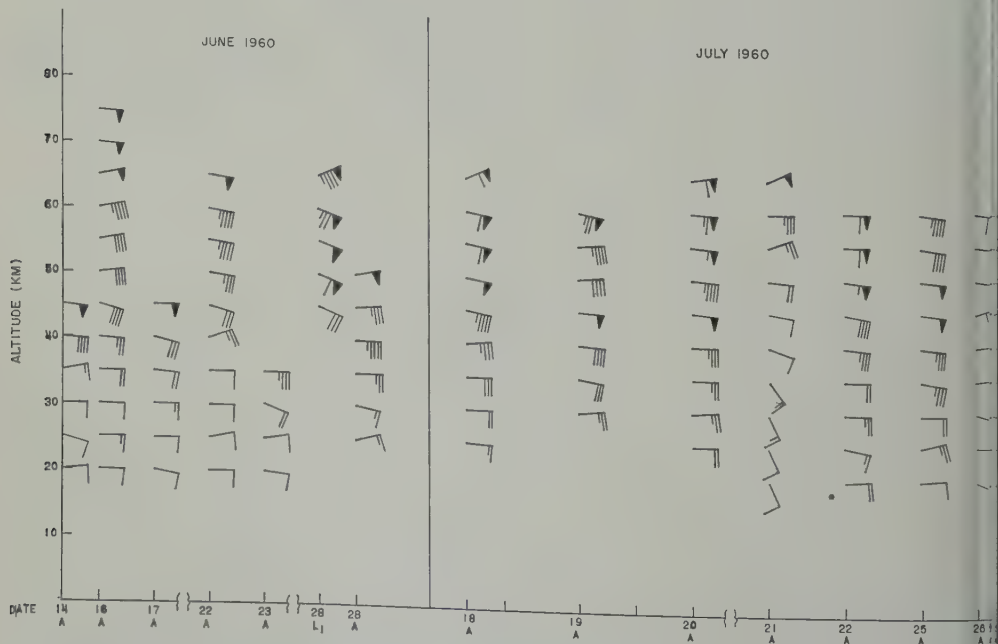


Fig. 11. Observed winds over White Sands, New Mexico, during June and July 1960. See Fig. 6 for explanation of symbols.

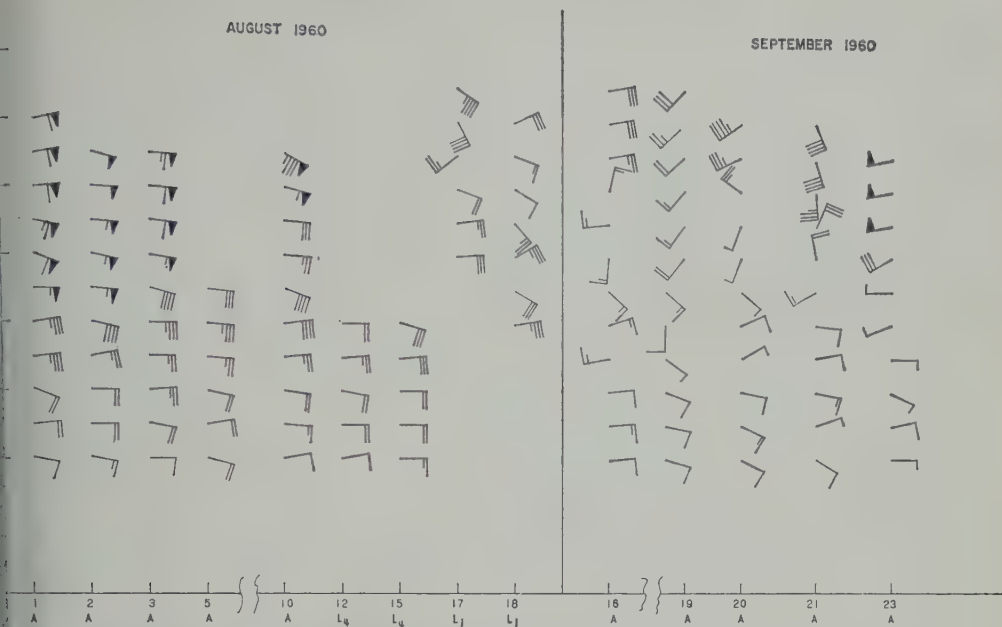


Fig. 12. Observed winds over White Sands, New Mexico, during August and September 1960. See Figure 6 for explanation of symbols.

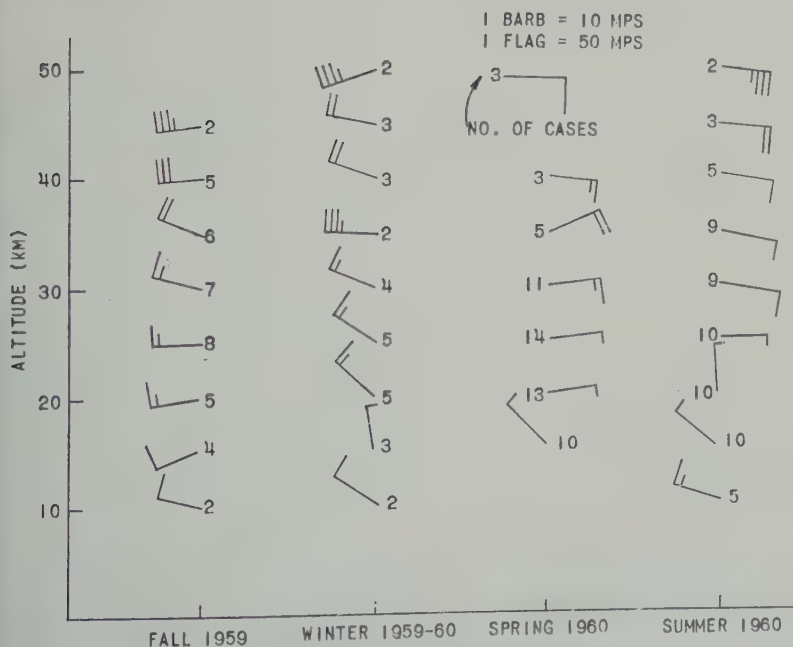


Fig. 13. Mean vector winds, Fort Churchill, Canada.

increase during the September observational period.

The mean wind profiles (Figs. 13, 14, 15, and 16) illustrate the seasonal characteristics of the wind field. They demonstrate quite clearly the monsoonal nature of the mesospheric circulation. In Figure 17 the monthly values of the N-S and E-W wind components at White Sands, New Mexico, are compared, and the rapidity of the circulation reversal in the mesosphere is dramatically illustrated.

Short-period variability of rocket winds. The time cross sections and seasonal profiles of wind velocity discussed in the previous section provide a qualitative idea of the wind variability with time over periods ranging from days to seasons. In this section we shall discuss the variability over the period of approximately 1 hour based upon data from Cape Canaveral, Florida, during the summer of 1960. The measure of variability in this analysis is

$$\Delta \bar{v} = \sum_{i=1}^n \Delta v_i$$

where ΔV_i is the change in the mean wind speed in a specified layer of the atmosphere between consecutive observations taken from 17 to 75

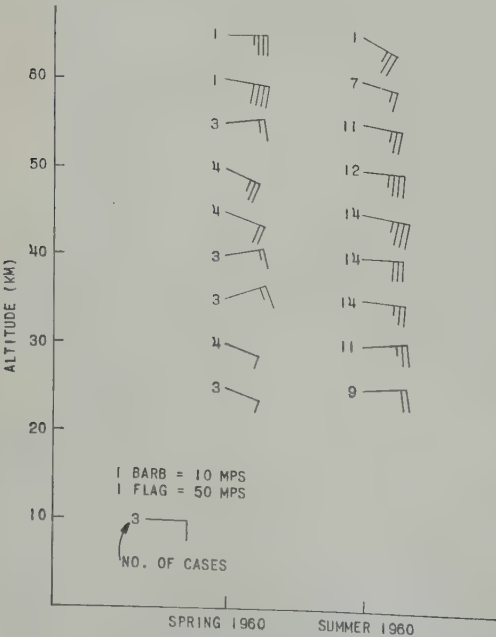


Fig. 14. Mean vector winds, Cape Canaveral, Florida.

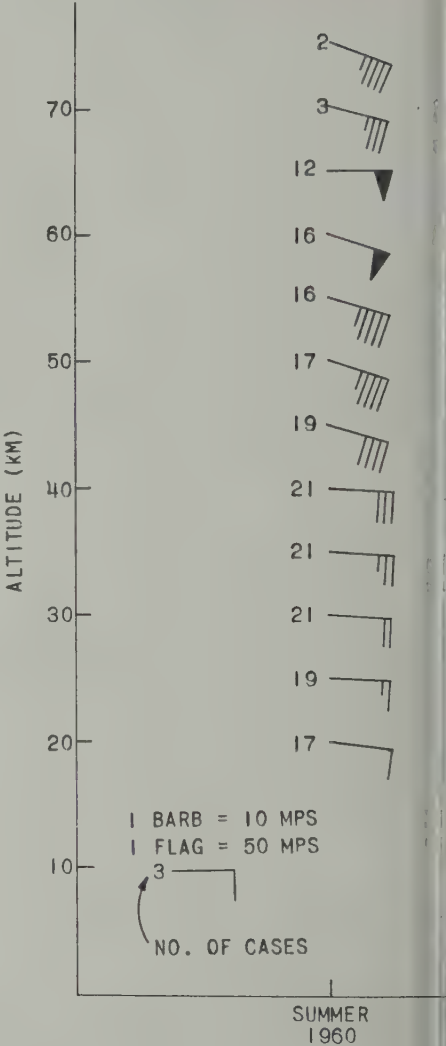


Fig. 15. Mean vector winds, White Sands, New Mexico.

minutes apart. The same procedure was used to determine wind direction.

There were nine days during July and August 1960 when two or more soundings were taken at .012-inch nylon chaff. On seven of these days there were two soundings; on one day there was one sounding; and on the remaining day, four. If all the collected data from all levels there would be a maximum number of 12 sets of wind data at each selected level from 25 km to 75 km. However, this was not the case. For example, at 65 km only six winds are available and only two are consecutive, giving only one chance to use for variability. This, of course,

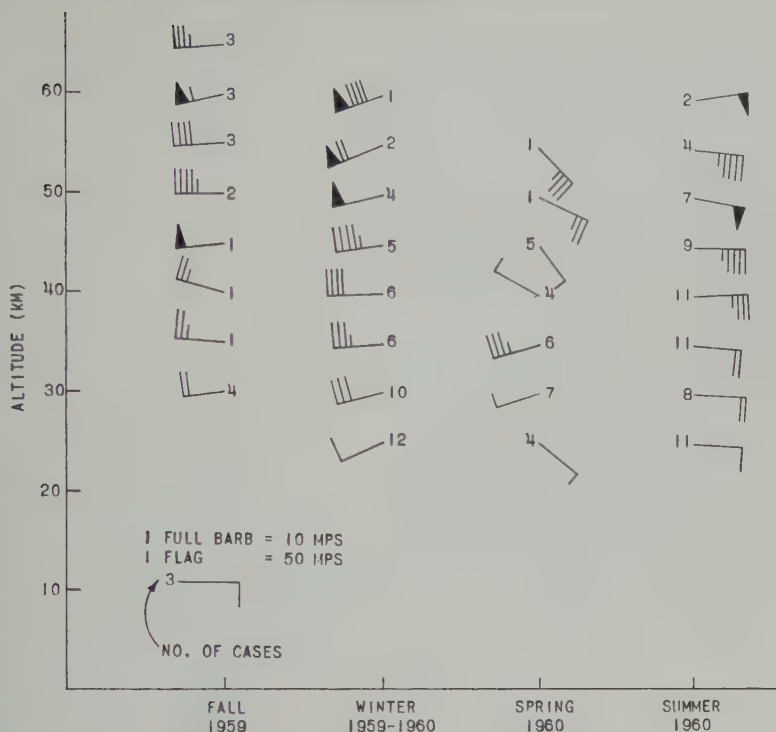


Fig. 16. Mean vector winds, Point Mugu, California.

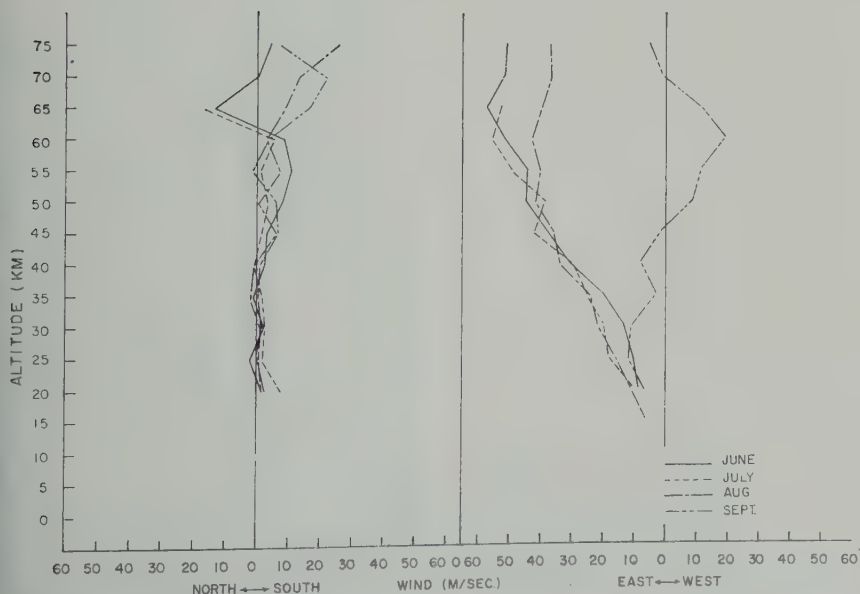


Fig. 17. Monthly mean wind velocity components, White Sands, New Mexico.

insufficient for any analysis. The data were therefore grouped into layers approximately 4.5 km thick to increase the reliability, and the resultant variability was plotted at the midpoint of the layer.

It can be seen from the curve in Figure 18 that the variability is dependent upon height, the mean difference in speed increasing from 2 m/sec at 25 km to 8 m/sec at 65 km. These values are somewhat higher than would be expected from studies of variability at lower levels, and the strong dependency upon height was not noticed in the low-level studies. It is in order to inquire to what degree the increasing variability with height is connected with increasing wind speeds. The average wind speed against height is plotted in Figure 19. It is plainly seen that the maximum wind is found near 50 km and that it tapers off sharply above that level. Clearly, the increased variability at the highest levels is not due only to increased

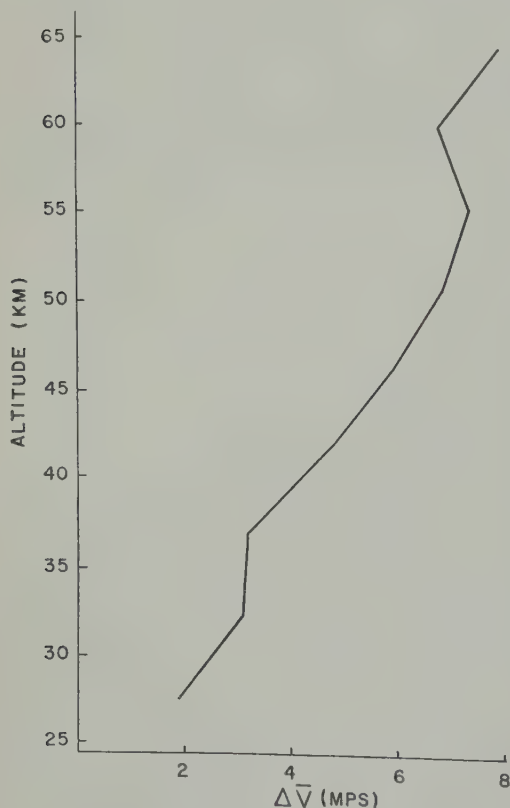


Fig. 18. Vertical distribution of short-period wind speed variability at Cape Canaveral, Florida, in summer.

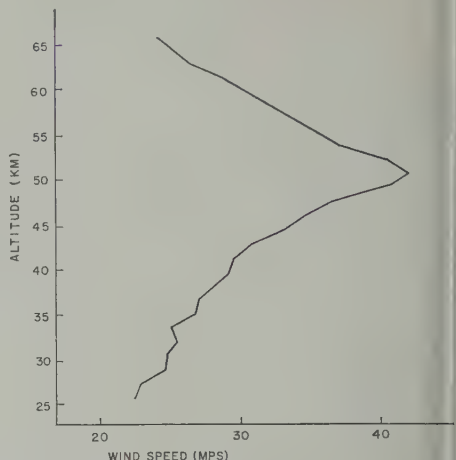


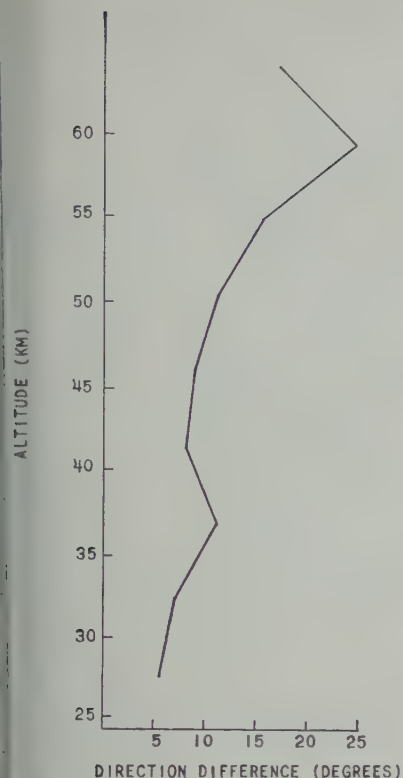
Fig. 19. Vertical distribution of wind speed at Cape Canaveral, Florida, in summer.

wind speeds, although the decreased slope of the variability curve above 50 km suggests that it is partially dependent on wind speed.

Another consideration is the increased variability in radar tracking at the higher levels. However, this does not apply to chaff (which was used in all the soundings in this study) because it initially presents a very good coherent target, but as it falls it disperses, causing the radar skip from point to point in a cloud that may cover a square kilometer in area. Hence, the variability due to errors in tracking the target actually decreases with altitude.

Tidal effects can be eliminated because the time increments are too small to include significant changes, if indeed there are notable effects at these levels.

It is also of interest to consider the variation of wind direction. Figure 20 shows the change in wind direction for 4.5-km layers. The direction is somewhat different from that of wind speed. There is little difference in direction variability from 25 km to almost 50 km, but large changes thereafter occur above 50 km. Figure 14 shows the winds during the period studied (August 8-25, 1960). From 25 to 45 km there is a persistent layer of easterly winds with a northerly component. It appears that Cape Canaveral lies on the eastern edge of a low pressure system to the north of the station. With increasing height above 45 km, the wind moves north of the station. Hence, a slight change in its position will change the wind direction from southeast to south or even southwesterly.



20. Vertical distribution of short period wind direction variability at Cape Canaveral, Florida, in 1960.

At 60 km the trough line is definitely northward station, as evidenced by the westerly wind components. A slight error in altitude will give a greater direction difference than would be at lower levels. It is believed that the direction differences at the highest elevations are attributable, in part, to this uncertainty in analysis. The potential of the data in synoptic and diagnostic investigations is demonstrated by the examples in this section. These examples consist of analyses of circulation over North America for selected months in the spring, summer, and fall of 1960. The flow was analyzed at 3-km intervals starting from 25 km. All available balloon and rocket data were used in the analysis; however, only the wind flow and that at a higher level are shown. To illustrate the variation in wind speed, a logarithmic wind scale was used to define the shape of the streamlines.

Figure 21 illustrates the flow at 25 km on September 17, 1960. A generally easterly circulation was established over the polar regions,

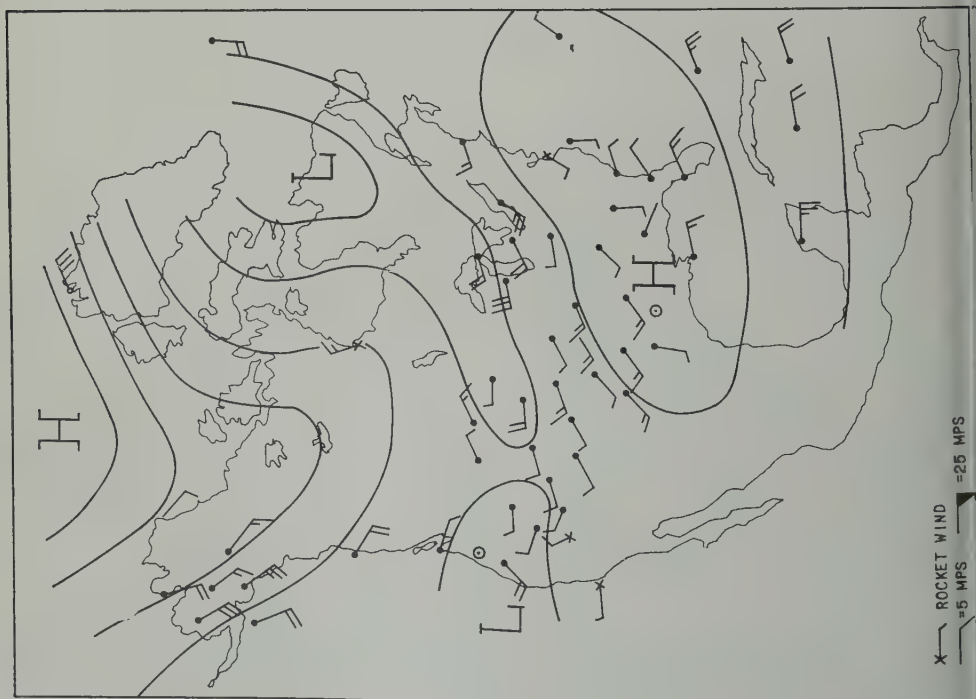
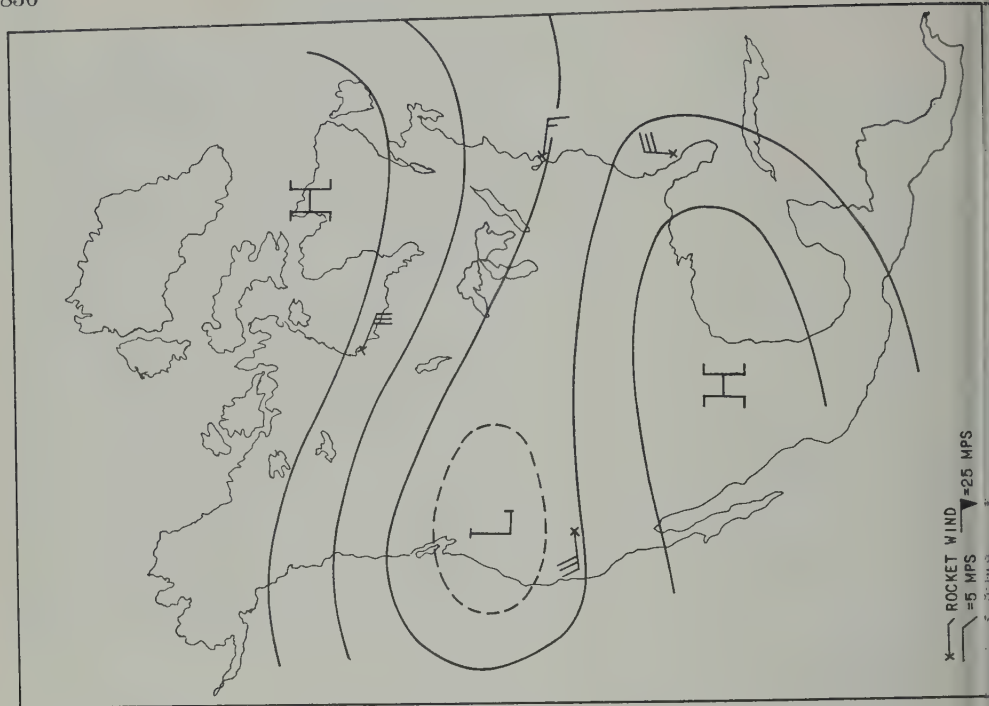
although disturbances in the zonal flow still exist, as evidenced by the trough extending northwestward over the Canadian archipelago from the cyclone over Labrador. This cyclone and its major trough, to the southwest, are remnants of the cold, polar-night vortex which has been displaced to lower latitudes because of the rapidly rising temperatures in the polar stratosphere. At 42 km (Fig. 22) considerable change has taken place. The easterly circulation is established more firmly over the Arctic. The cyclone over Labrador has been completely filled by the warm polar air and has been replaced by a ridge of the polar anticyclone. Examination of the wind profiles at Fort Churchill and Wallops Island reveals that there was a smooth transition from the cyclone at 25 km to the ridge at 42 km. At Churchill the wind veered steadily from NW to SE while at Wallops Island the wind veered slightly from SW at 25 km to NW at 35 km and then sharply to SE above 35 km as the trough line retreated southward.

The mid-latitude cold belt responsible for the trough which lay along the Canadian-United States border at 25 km is no longer present at 42 km. The only vestige of it is a zone of cyclonic shear across the United States which weakens with elevation above 35 km. Examination of data to higher levels on surrounding days indicates that between 45 and 50 km the high- and mid-latitudes of North America were completely dominated by easterly flow.

The summertime flow at 25 km (Fig. 23) displays the easterly circumpolar current which is characteristic of the stratosphere in this season. The only apparent change with elevation is a steady increase in the speed of the wind (Fig. 24). Examination of day-to-day variation in the wind velocity suggests that disturbances are present in the easterly flow, but they cannot be resolved by this small network.

Figure 25 is a 1-mb map for the 1960 summer season. The contours are based upon geostrophic winds. Heights were computed from average summer temperature data over White Sands referred to a 15-mb datum level.

In the early fall there is again a considerable change in the circulation from the stratosphere to the middle mesosphere. The 25-km chart for September 20, 1960 (Fig. 26), shows that at high latitudes the winter westerly circulation



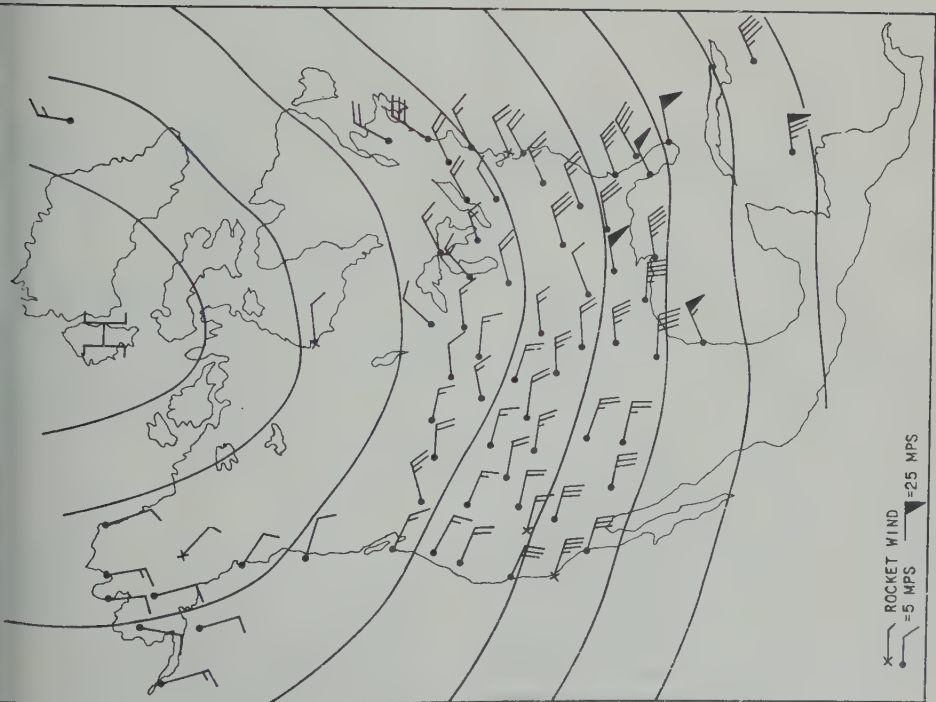


Fig. 23. Circulation at 25 km, July 20, 1960.

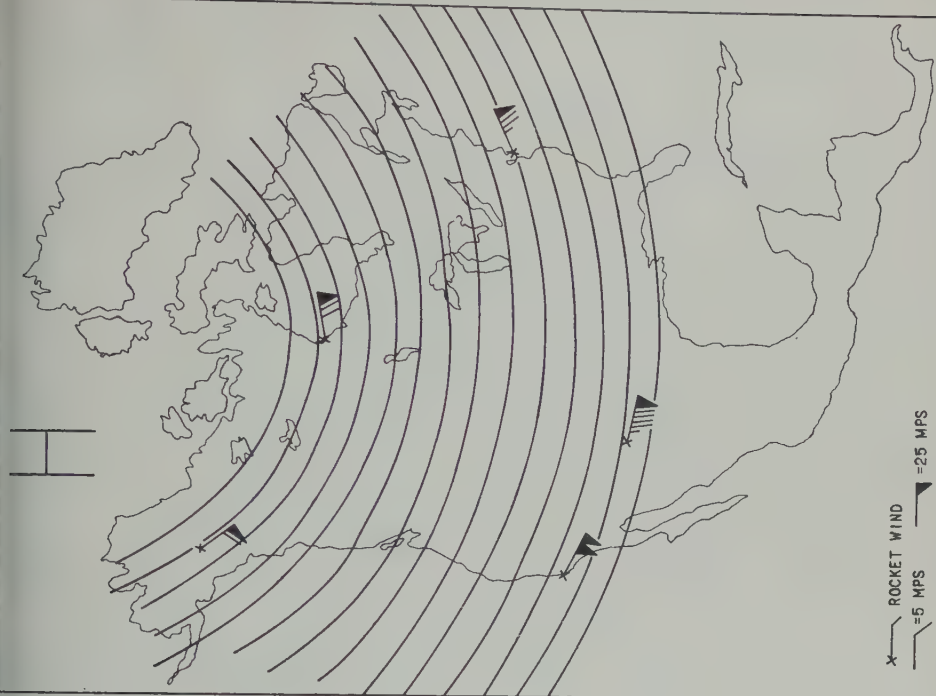


Fig. 24. Circulation at 50 km, July 20, 1960.

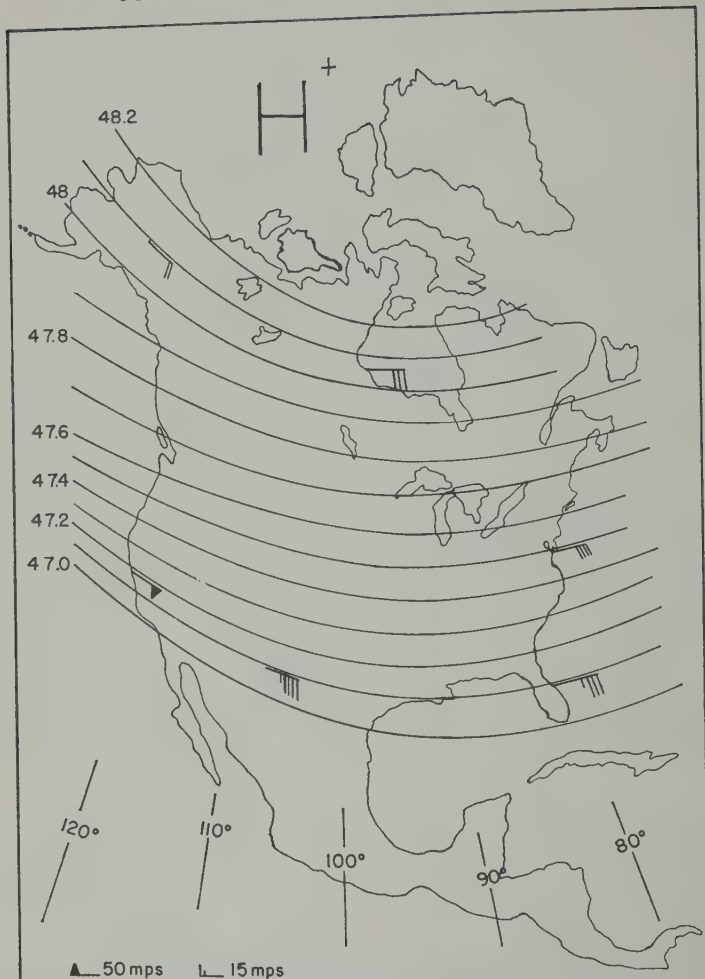


Fig. 25. Circulation at 1 mb, summer 1960. Height contours are in geopotential kilometers.

has been established. The ridge, or anticyclonic shear zone, across the United States is indicative of a mid-latitude warm belt between the colder polar and equatorial stratospheres. Between 25 and 43 km the location of the shear zone remained essentially unchanged, but westerly winds increased steadily with elevation to the north while the easterly winds to the south remained steady. At 50 km (Fig. 27) the winds over the southwestern United States have shifted from easterly to westerly components and a trough extends along all of western North America. The warm belt no longer existed over the area covered by the analysis, and only a weak remnant of the anticyclonic shear zone remained over the southeastern United States.

A winter 1-mb map is shown in Figure 28. It was constructed in the same manner as the

summer chart in Figure 25 except that temperature data were used.

Although it is risky to base generalizations upon a few selected cases, this series of maps suggests that seasonal changes in the circulation originate at high latitudes above the mesosphere and work their way downward and southward. This feature can be further observed by examination of the time cross sections of winds (Figs. 13 through 16).

SUMMARY AND CONCLUSIONS

The experience gained in collecting and analyzing the data during the first year of the Meteorological Rocket Network leads to the following conclusions about the feasibility of a network and the value of the data.

There is no doubt that a network can

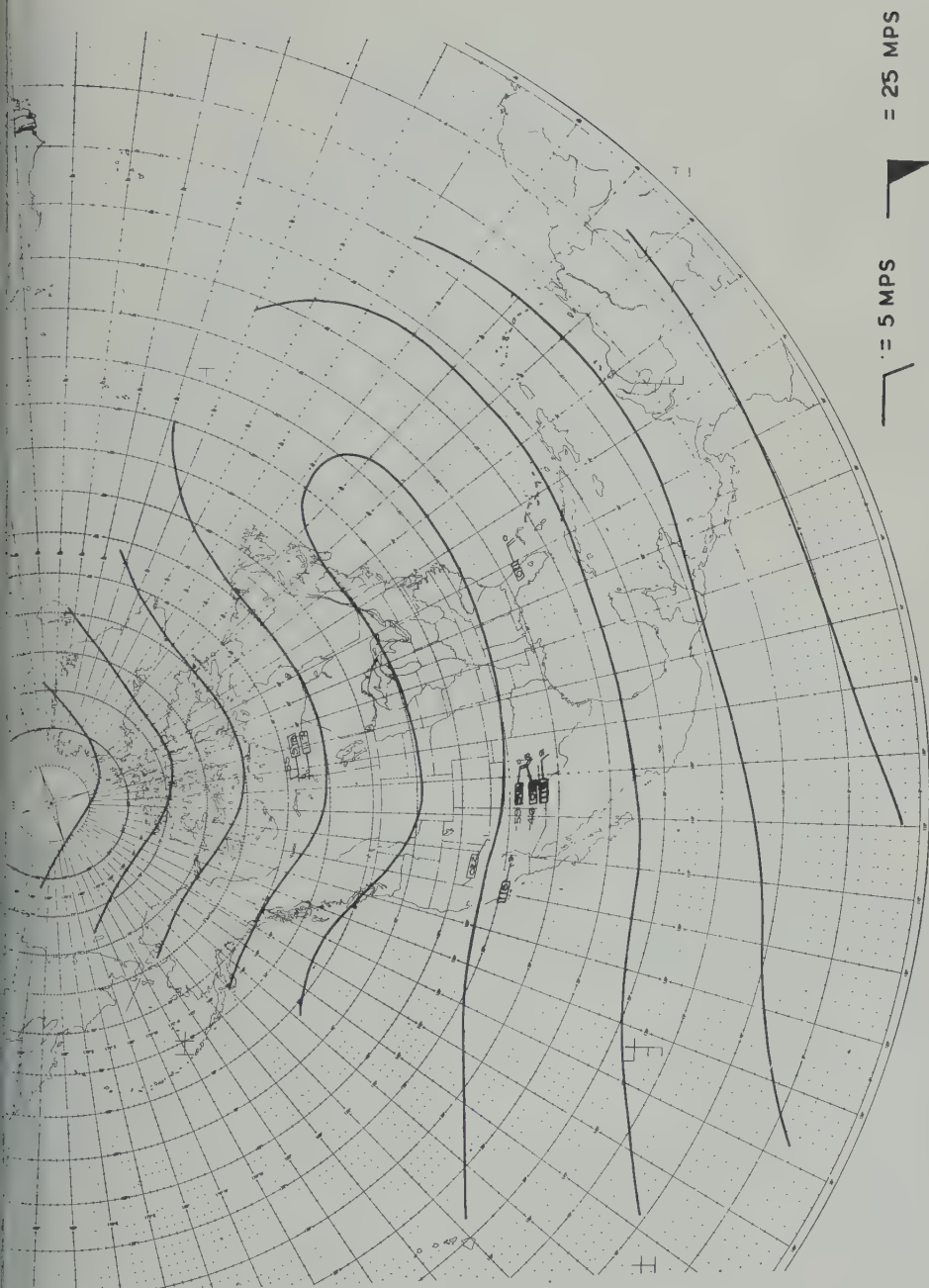
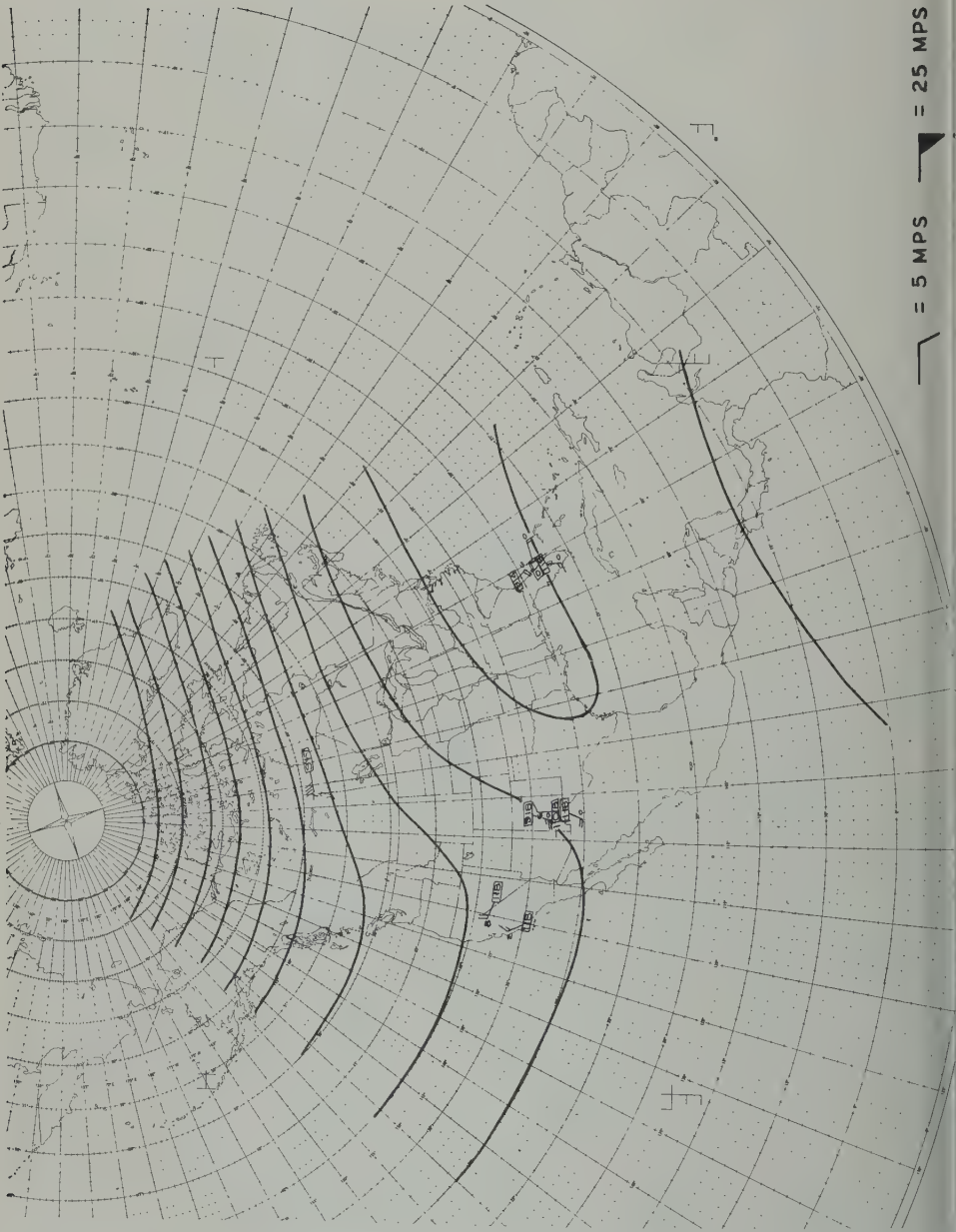


Fig. 26. Circulation at 25 km, September 20, 1960. (Balloon winds were used in making analysis but were not plotted on chart.)



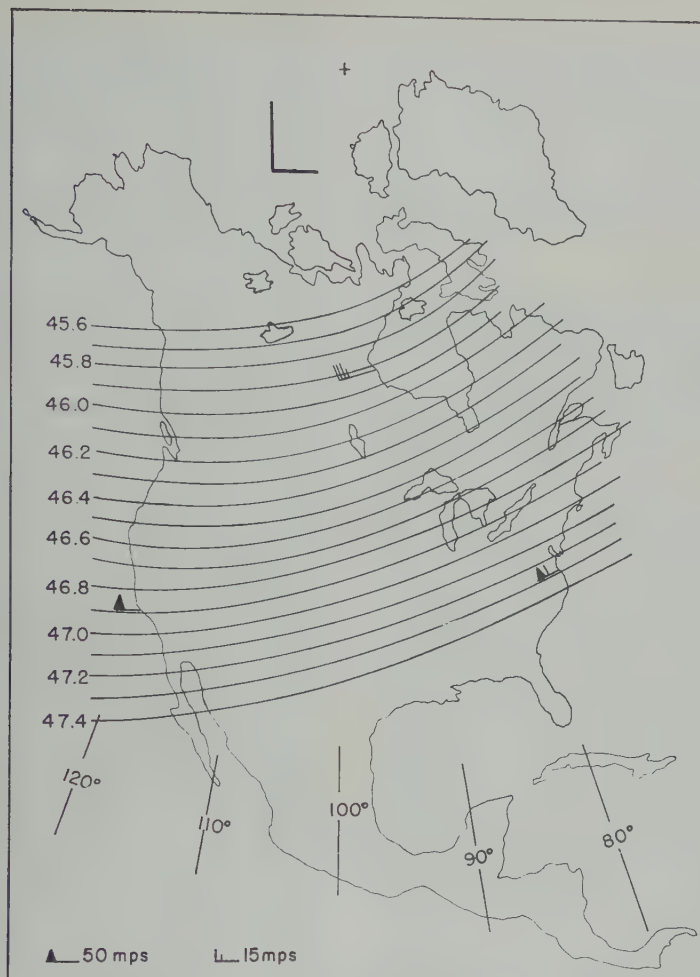


Fig. 28. Circulation at 1 mb, winter 1959-1960. Height contours are in geopotential kilometers.

fairly rigid predetermined schedule, provided the schedule is compatible with the limits of the ranges. In the future it may even be possible to achieve a certain amount of independence from the missile ranges. The Meteorological Rocket Network has made available for the first time, significant amounts of meteorological data from the previously 'silent' region between the top of the radiosonde layer and the bottom of the ionized layer. As important, perhaps, as the quantity of the data is the ease of reduction. Sophisticated techniques which automatic computers are used should be applied to get the full amount of information available in the sounding, but wind information of sufficient accuracy for many purposes can be obtained graphically in less than an hour. This

makes the technique suitable for field use.

Scientific results of the initial analysis of the data include an indication of how the monsoonal circulation reversals proceed. In both spring and fall the shift from westerlies to easterlies and back again appears to occur first at high latitude and above 50 km. As the season advances the change propagates downward and equatorward. The data also suggest that the circulation in mid- and upper-mesosphere responds more rapidly than that of the lower levels to the seasonal variation in solar energy received by the atmosphere.

The short-period variability of wind speed has been shown to increase with altitude. Although there is no immediately available explanation of this behavior, it is consistent with

the strong vertical wind shears observed above the mesopeak during the same series of observations. It would also be expected that the short-period variability or turbulence would be greater above the mesopeak where the temperature decreases with height as it does in the troposphere. This of course does not explain why the variability increases in a near linear fashion throughout the mesocline.

Temperature measurements were scarce during this first year, but it was gratifying to find that the available data up to 50 km agreed in the mean with the values in the standard atmosphere.

Circulation analyses at levels up to 50 km have been included in this report. Despite the wide spacing of stations, these charts were valuable in making diagnostic studies. The continuous improvements in equipment and techniques assure us that the analyses will be carried out to higher levels in the future.

Acknowledgments. We wish to acknowledge the contributions of Dr. Hans aufm Kampe of the U. S. Army Signal Research and Development Laboratory and Mr. Willis Webb of the U. S. Army Signal Missile Support Agency, whose individual efforts are largely responsible for the initiation of the Meteorological Rocket Network. Our appreciation is also extended to all those at the missile ranges who helped make the program a success, particularly Lcdr. W. E. Hubert, Point Mugu; Mr. Kenneth R. Jenkins, White Sands; Maj. Robert Miller, Cape Canaveral; Mr. Lawrence Smith, Tonopah; Mr. John Spurling, Wallops Island; and Mr. Otto Thiele, Fort Churchill.

REFERENCES

Ballard, H., Response time of and effects of radiation on Veco bead thermistor, Schellenger Res.

- Lab., Texas Western College, *USASMSA DA 29-040-ORD-2140*, El Paso, Texas, 1961.
- Barr, W., Theoretical evaluation of cylindrical as a wind sensor at high altitude, *USASMSA Tech. Rept. 2138*, U. S. Army Signal Research and Development Laboratory, Fort Monmouth, N. J., 1960.
- Barr, W., Theoretical considerations in the use of atmospheric temperature sensing elements, *USASRD Tech. Rept. 2195*, U. S. Army Signal Research and Development Laboratory, Fort Monmouth, N. J., 1961.
- Clark, G., Development of a rocket telemetry package for the Meteorological Rocket Network, published as an *IRIG Document*. 1961.
- Jenkins, K., and W. Webb, High altitude wind measurement, *J. Meteorol.*, **16**, 511-515, 1959.
- aufm Kampe, H., The Meteorological Rocket Network, *Weatherwise*, **13**, 192-195, 1960.
- Rapp, R., The accuracy of winds derived from radar tracking of chaff at high altitudes, *J. Meteorol.*, **17**, 507-514, 1960.
- Smith, L., The measurement of winds between 100,000 and 300,000 ft. by use of chaff rockets, *J. Meteorol.*, **17**, 296-310, 1960.
- Wagner, N., Theoretical time constant and radiation error of rocketsonde thermistor, Schellenger Res. Lab., Texas Western College, *USASMSA contract DA 23-072-ORD-1564*, El Paso, Tex., to be published, 1961.
- Webb, W., W. Hubert, R. Miller, and J. Spurling, Initiation of the Meteorological Rocket Network, presented at 187th Meeting, Am. Meteorological Society, Eugene, Oregon, 14-16 June, 1960.
- Webster, R., W. Roberts, Jr., and E. D. Schellenger, Development of the Arcas rocketsonde system, *Final Rept.*, contract *NONr-2477(00)*, Aerospace Research Corp., Alexandria, Va., 1960.

(Manuscript received May 28, 1961.)

Experimental Flight Verification of the Economical Net Radiometer

R. H. BUSHNELL¹ AND V. E. SUOMI

*Department of Meteorology, University of Wisconsin
Madison, Wisconsin*

Abstract. In an experiment to verify the effect of convection and air conduction on the economical net radiometer, a flux plate type of radiometer, along with its required d-c amplifier, was designed, built, and flown with a radiosonde. Two such flights were made. The net radiation values given by the flux plate agreed well with the values given by the economical net radiometer, showing that convection does not occur in nighttime use of the economical radiometer and that the air conductivity function being used is valid. An estimate is given of the precision of radiative cooling values which can be obtained from economical radiometer data.

Introduction. Two test flights of balloon-radiometers have been made in an attempt to obtain *in situ* system tests of the simple radiometer developed at the University of Wisconsin. This test is part of a continuing study to determine as completely as possible the performance of the instrument under all flight conditions.

This instrument, described in detail elsewhere (Suomi and Kuhn, 1958; Suomi, Staley, and Kuhn, 1959), is easily attached to the U. S. radiosonde. A considerable number of radiometer flights have now been made owing to the cooperation and help of the U. S. Weather Bureau. The usefulness of radiometer soundings is that they make possible measurements of the real derivative of the net radiation. This radiometer is directly related to the cooling or heating of the atmosphere by long-wave radiation. Areal differences in cooling rate can be expressed as power available to drive the atmosphere.

The upward and downward streams of radiation are calculated from thermistor-measured temperatures of the sensor surfaces and air temperature. Net radiation is calculated from the difference of these temperatures. The rate of heating or cooling of the atmosphere, on the other hand, is obtained by observing the change in temperature difference with height. Clearly the heating or cooling of the atmosphere is derived from a temperature difference of the sensor temperatures. Although the percentage error of each sensor

temperature may be low, it is easy to show that the percentage error of the difference can be higher, the degree of the degradation being inversely related to the magnitude of the difference. In a typical flight of the radiometersonde the difference between temperatures of the sensors near the earth's surface is about 10° , and it increases to $25\text{--}30^{\circ}$ at 30 km. Now the size of the second difference and hence its percentage error depends on the size of the pressure interval over which the second difference is taken. Thus, although the heating or cooling for a deep layer in the atmosphere can be measured to good accuracy, the accuracy is not good for a thin layer of say 10 mb. Clearly, the real test of the instrument is in the stratosphere where both the pressure intervals and the changes of radiation with height are small. If one is interested in studying departures from radiative equilibrium in the stratosphere, cooling rates as low as 0.1° per day could be important [Goody, 1950]. Our results show that for a 100-mb pressure interval the standard error of the present radiometersonde system in heating or cooling rates is $0.15^{\circ}/\text{day}$. The important point is that the limitation on accuracy is not at present determined by the radiometer instrument but by the radiosonde telemetry system. The temperature sensitivity of the system becomes low at low temperatures. We have subsequently used a simple modification of the ground equipment whose purpose is to increase the low-temperature scale factor and have reduced the uncertainty by a factor of 2. Therefore it is possible to determine cooling rates of $.08^{\circ}/\text{day}/100\text{ mb}$.

When one considers the heating or cooling of deep layers in the atmosphere, such as the troposphere or the entire atmosphere, the precision of measurement increases accordingly. It is easy to detect a cooling rate difference of only .02°/day. Cooling rates of 1.5°/day for arctic air and 3°/day for tropical air are typical, so accuracies of this measurement of better than 2 to 3 per cent are possible.

A table in which the magnitudes of the errors of other instruments are listed for comparison is included. These values were taken as those either mentioned or indicated in published figures or descriptions and in most cases are only mentioned incidentally and not as a result of a published error study. The reader is cautioned to keep this in mind.

Within the limitations outlined above, soundings of net radiation to good accuracy are possible and appear to be a very useful addition to the air temperature, pressure, and humidity measurements obtained in a typical radiosonde sounding.

Principle. Although the economical net radiometer has been shown to be correct on the ground [Tanner, Businger, and Kuhn, 1960; Kuhn, 1961], some question has remained about its behavior high in the atmosphere. In particular, it was assumed that, because the Grashof number [de Graaf and van der Held, 1953] was always less than 1000, there was no convection in the thin air cells that protect the absorbing surfaces from ventilation. Also, the conductivity of the air in these cells was taken to be a certain function of temperature. These assumptions had not been verified experimentally for the radiometer in flight with its wide range of pressure and temperature.

Further, it has been suggested [Houghton, 1958; Gergen, 1958] that because the calculation of net radiation involves the difference of two numbers, the economical net radiometer might suffer from large errors which would make it almost useless for finding radiative cooling. Although Kuhn, Suomi, and Darkow [1959] indicate that this is not the case, they do not give experimental data.

Accordingly, we have done the experiment reported here, consisting of two radiosonde flights each carrying an economical net radiometer and a flux plate radiometer. We adapted a ventilated flux plate radiometer [Suomi,

Franssila, and Islitzer, 1954] to balloon flight did not use ventilation because of the required and because ventilation would with air density anyway. Instead, we used polyethylene convection shields, the same fact as we have used on the economical radiometer. The flux plate radiometer is the same as the economical radiometer except it has two blackened surfaces and polyethylene convection shields. In both, radiant energy is absorbed and then disposed of by reradiation and conduction. The difference in the two is in the relative amounts of reradiation and conduction. In the flux plate radiometer type, 99.5 per cent of the net radiation energy is conducted through the glass flux plate, only 0.5 per cent is conducted away through the air, and 0.3 per cent is reradiated. In contrast, in the economical radiometer, approximately 50 per cent is conducted through the air cells and 50 per cent is reradiated. With this extreme difference, the two radiometers, on the principle, should still agree in measured flux of radiation. This agreement is the verification sought in the experiment.

The flux plate develops a maximum temperature difference from top to bottom of 0.3°C, so that large temperature gradients do occur in the adjacent air cells formed by the shields. The data from the flights show that, at times, the temperature gradients were as high as zero. Zero or not, we are assured that there is no significant air conduction and that the Grashof number is far smaller in these cells than in the economical radiometer.

The flux plate gives a recorder deflection essentially proportional to net radiation, and, if it, there is no problem of the difference of two numbers.

In addition, as the flux plate radiometer has a time constant of 0.8 second, by using it to check the corrections made for the 400 second time constant of the economical radiometer.

Structure of the sondes. The flux plate consists of 68 turns of constantan wire wound on microscope slides and copper plated thermopiles, were painted black with the kind of paint used on the economical radiometers.

For the first one flown, the polyethylene convection shield consisted of four layers of $\frac{1}{4}$ inch apart. The flux plate was n



1. Flux plate radiometer used on the second flight.

horizontally in the center cell. As polyethylene is transparent, one could see through the radiator structure everywhere except through the cells themselves and through the wood supports. In this arrangement radiation entering one side of the radiometer could pass through to the other side and be reflected by the polyethylene to the wrong face of the flux plate—a distinct disadvantage. After finding on the first flight that the temperature gradients in the air cells were indeed as small as we had expected, we changed the design for the second radiometer in two ways. We used only one polyethylene sheet on each side of the radiometer mounting, and we put a black, opaque film in the plane of the flux plate so that no radiation could reach the wrong side of the plate. This radiometer is shown in Figure 1.

The 0- to 700-microvolt signal from the flux plate was telemetered to the ground and recorded on a standard radiosonde receiver. In this radiometer the flux plate signal was amplified by a chopper type d-c amplifier whose output controlled the frequency of the radiosonde timing oscillator. The system was similar to that described by *Suomi and Barrett* [1952]. The flux plate signal, along with a reference voltage from a mercury cell, flux plate temperature, air temperature, temperatures from the chemical radiometer, and monitoring signals, were selected in sequence by a motor-driven switch once every 30 seconds during flight.

The complete radiometersonde, consisting of the radiosonde, an amplifier, two radiometers, and batteries, weighed 4200 grams. The flux plate radiometer itself weighed only 35 grams.

Because the thermal contact between the thermocouples and the glass plate is question-

able, the flux plate scale factor cannot be obtained from its dimensions. It must be found by calibration. On the other hand, if the assumptions are correct, the scale factor of the economical radiometer is known without comparison with a standard instrument. In view of this, we have made the comparison between the two radiometers by calculating the scale factor of the flux plate for each 30 seconds of flight, using the economical radiometer as the standard. If the assumptions are not correct, this scale factor will, as a consequence, change with height.

The following items were included in the calculation of net radiation from the telemeter record of flux plate signals: response produced by the reference voltage, linearity correction to the telemeter, change of thermocouple emf with temperature of the flux plate (17 per cent change during a flight), change of thermal conductivity of glass with temperature, change of conductivity of air with temperature. The transmissivity and reflectivity of polyethylene and the absorptivity of the black paint, which are known for ground level conditions [*Suomi, Staley, and Kuhn*, 1958; *Tanner, Businger, and Kuhn*, 1960], were assumed not to change with height. For both radiometers the conductivity of air was taken to be $(348.4 + 1.06T)10^{-5}$ cal/(cm min °C) where T is the mean Celsius temperature of the air in a cell.

It should be noted that, to ensure having reliable data from the flux-plate-amplifier system, we used four of the nine segments on the motor-driven switch for monitoring. One segment connected the reference voltage to the input of the amplifier, another put a short circuit on the input, the third gave the temperature in the amplifier box, and the fourth gave the filament voltage. We used these signals in accounting for voltage drift, in addition to using them to observe the condition of the sonde.

As the economical radiometer for use with radiosondes contains standard U. S. Weather Bureau thermistors, we telemetered its temperatures in the conventional manner, just like air temperature.

Flights and results. We made two night flights, both launched by the U. S. Weather Bureau radiosonde station at Green Bay, Wisconsin. In Figure 2, which shows the results for both flights, each point comes from a separate reading of the record; no smoothing has been

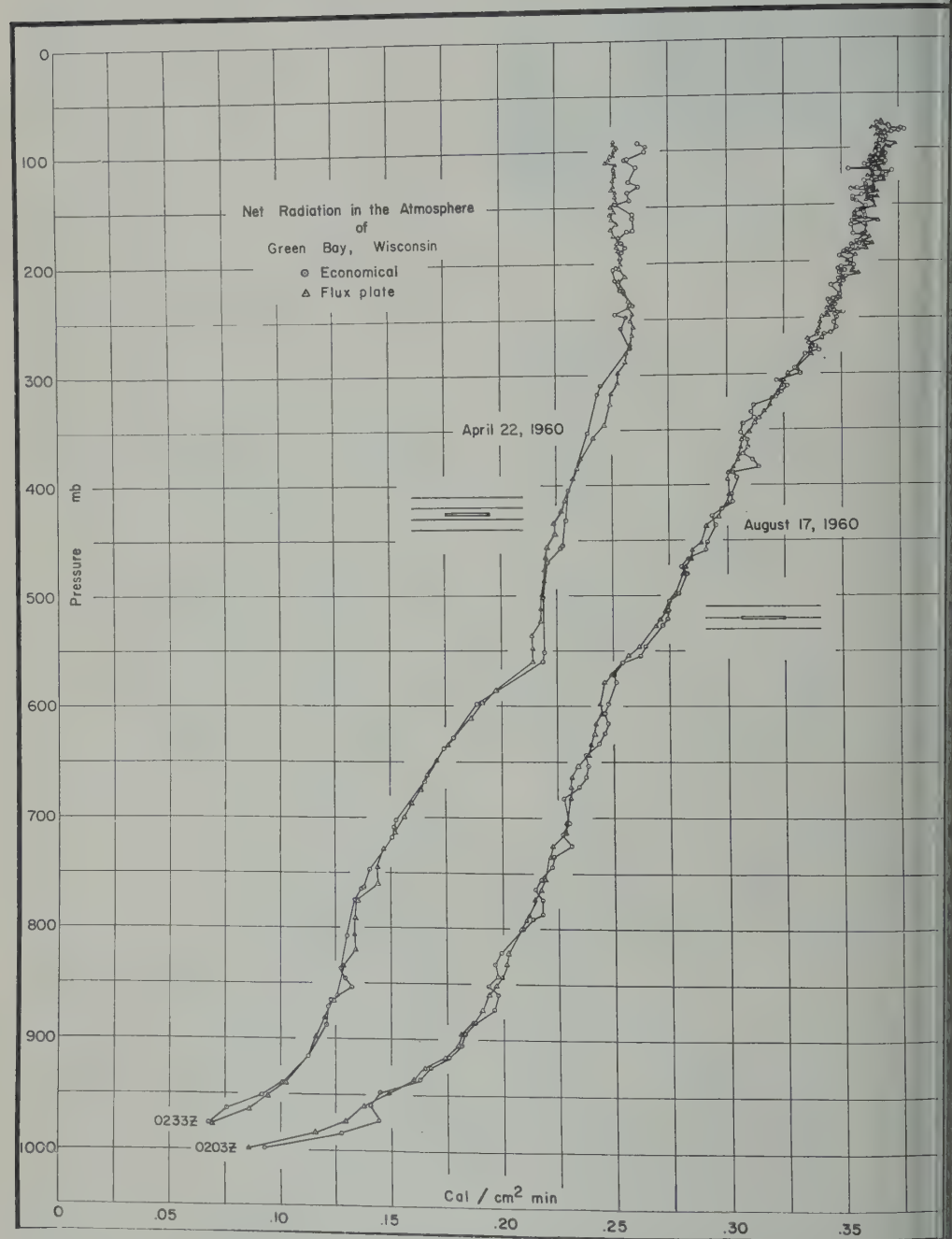


Fig. 2. Comparison of net radiation measurements made by the economical radiometer and the flux plate radiometer.

TABLE 1. Precision of Radiative Cooling Values

	Standard Error, °C/day	95% of Values within: (°C/day)
smoothing		
50 mb	0.6	1.2
100 mb	.3	.6
smoothed		
50 mb	.3	.6
100 mb	.15	.3

ed. The values shown for the flux plate obtained by using the mean scale factor the whole flight. As can be seen, no large discrepancies appeared. However at the top of April 22 flight the flux plate value is about 1 per cent below the economical radiometer value. This difference, caused by reflection from polyethylene in the manner discussed above, does not appear in the second radiometer having opaque film in the middle.

The fluctuations of the economical radiometer values are more or less the same over the whole flight and are larger than those of the flux plate radiometer. They show the effects of the difference of two temperatures and of the error in compensating for the long time constant. For the August 16 flight the rms fluctuation of the scale factor was 1.4 per cent.

Estimate of the errors of cooling values. From the data we can estimate the error of radiative cooling values obtained from economical radiometers. First we estimate the standard error of economical radiometer net radiation values (Table 1). The difference of the two radiation values at any one time in a flight is the combined error of both radiometers. This difference has a standard deviation of .0044 cal/(cm²min). As the errors of the two radiometers are independent, the variance of this difference is the sum of the variance for the flux plate radiometer and the variance for the economical radiometer. As a conservative estimate, based on the fact that the variance of the economical radiometer values is twice that of the flux plate radiometer, the standard error for the economical radiometer is then 0.817 times the combined error. Thus we estimate the standard error of a radiation value from an economical radiometer to be approximately .0036 cal/(cm²min), including the errors introduced by the calculations used to compensate for the time constant.

The rate of radiative temperature change is given by $\partial T/\partial t = -(g/c_p)\Delta R_N/\Delta p$. The precision of $\partial T/\partial t$ is proportional to the precision of ΔR_N and inversely proportional to Δp . ΔR_N , the difference of R_N at the top of a layer and R_N at the bottom, involves two measurements of R_N , but, because their errors are independent, the standard error of ΔR_N is $\sqrt{2}$ times the standard error of net radiation.

To illustrate these errors, we consider two cases, a 50-mb layer and a 100-mb layer. For these two we find the errors of $\partial T/\partial t$ given in the first two lines of Table 1. In getting these values no use has been made of any kind of smoothing. Much better results can be had by combining the information in adjacent measurements in a flight. For example, if numerical smoothing is applied to the net radiation data so that the information from four points yields one smoothed point, then, because the four measurements are independent, the error of the cooling values is cut in half, as is shown in the last two lines of Table 1. This precision, twice as good as that given by *Kuhn, Suomi, and Darkow* [1959], appears to be routinely obtainable in nighttime soundings if the record is carefully read.

Conclusion. No gross differences appeared between the two radiometers. The differences which did appear are small and random. Therefore, within the accuracy of these measurements, convection does not occur in nighttime use of the economical radiometer and the air conductivity function given above is valid. Radiation values obtained from the economical radiometer are not biased by large changes with height but are contaminated mainly by random fluctuations. The necessity of using the difference of two numbers in reducing economical radiometer data does not appear to cause serious errors. In spite of its fast response, the flux plate gave a smoother sounding than the economical radiometer gave. From this we conclude

TABLE 2. Precision of Net Radiation Values

Measurement	95% of Values within: (cal/cm ² min)
<i>Pohl</i> [1956]	0.03
<i>Aagard</i> [1960]	.04
<i>Fenn and Weickmann</i> [1960]	.03
Economical radiometer	.007

that the long time constant of the economical radiometer does not cause radiation structure to be missed.

The soundings of Pohl [1956], Aagard [1960] and Fenn and Weickmann [1960] are different from ours, showing more and larger changes between heating and cooling. The comparison of the radiometers in this experiment shows that the economical radiometer would have indicated such extremes if they had existed at the times of our flights. Most other soundings made by economical radiometers likewise do not show such a number of changes between layers of large heating and cooling. Although the instruments have not been experimentally compared, it would seem that the differences are a matter of the accuracy of the measuring systems and the smoothing used. In this connection, Table 2 shows the precision of the measurements reported by these authors.

Acknowledgments. I wish to thank the staff of the U. S. Weather Bureau Airport Station at Green Bay, Wisconsin, for conducting the two soundings. P. M. Kuhn assisted in this experiment, which was supported by the Office of Meteorological Research, U.S. Weather Bureau.

REFERENCES

- Aagard, R. L., Measurements of infrared radiation divergence in the atmosphere with the double-radiometer and the black ball, *J. Meteorol.*, **17**, 311, 1960.
de Graaf, J. G. A., and E. F. M. van der Held, The

- relation between the heat transfer and the convection phenomena in enclosed plane air layer, *Appl. Sci. Research*, **3 A**, 393, 1953.
Fenn, R. W., and H. K. Weickmann, Atmospheric net radiation flux during winter in the Thule area, Greenland, *J. Geophys. Research*, **65**, 3651, 1960.
Gergen, J. L., *J. Meteorol.* **15**, 412, 1958.
Goody, R. M., The experimental determination of the thermal state of the lower stratosphere, *Centenary Proc. Roy. Meteorol. Soc.*, p. 9, 1958.
Houghton, J. T., *Quart. J. Roy. Meteorol. Soc.* **84**, 472, 1958.
Kuhn, P. M., V. E. Suomi, and G. L. Darr, Soundings of terrestrial radiation flux over Wisconsin, *Monthly Weather Rev.*, **87**, 129, 1959.
Kuhn, P. M., Accuracy of the airborne economical radiometer, *Monthly Weather Rev.*, August 1960.
Pohl, W., Messungen des Ultraroten Strahlungsstromes in der freien Atmosphäre, *Z. Geophysik*, **22**, 1, 1956.
Suomi, V. E., and E. W. Barrett, An experimental radiosonde for the investigation of the distribution of water vapor in the stratosphere, *J. Sci. Instr.*, **23**, 272, 1952.
Suomi, V. E., M. Franssila, and N. F. Islitzen, Improved net radiation instrument, *J. Meteorol.*, **11**, 276, 1954.
Suomi, V. E., and P. M. Kuhn, An economical radiometer, *Tellus*, **10**, 160, 1958.
Suomi, V. E., D. O. Staley, and P. M. Kuhn, Direct measurement of infra-red radiative divergence to 160 mb, *Quart. J. Roy. Meteorol. Soc.* **84**, 134, 1958.
Tanner, C. B., J. A. Businger, and P. M. Kuhn, The economical net radiometer, *J. Geophys. Research*, **65**, 3657, 1960.

(Manuscript received March 30, 1961; revised May 18, 1961.)

A Method of Obtaining Detailed Wind Shear Measurements for Application to Dynamic Response Problems of Missile Systems

HAROLD B. TOLEFSON AND ROBERT M. HENRY

*NASA Langley Research Center
Langley Field, Virginia*

Abstract. The development of efficient missile systems requires more precise measurements of the winds and wind shears within the atmosphere than are provided by current wind-sounding methods. The gross smoothing inherent in balloon soundings effectively filters out the small-scale wind variations which are of considerable importance to missile response problems. A method in which the detailed structure of the vertical wind profile is determined from photogrammetric measurements of the successive positions of a smoke trail left by a vertically rising missile has been developed and applied by the Langley Research Center of the NASA. The method is described in this paper and its accuracy is discussed. Results in the form of wind profiles for several measurements are presented. Particular attention is given to the continuous and random nature of the wind disturbances as a function of altitude and to the implication of these disturbances characteristic to missile responses.

INTRODUCTION

The fluctuating winds experienced by a missile during flight through the atmosphere represent one of the important dynamic loading effects for missiles. Rapid changes in wind velocities with altitude, or wind shears, are of importance not only because of the structural stresses induced but also because possible excitation of structural vibration modes by the unsteady dynamic loading conditions may affect the performance of the guidance and control system. In many missile trajectories, the most critical loading occurs near tropopause levels (about 30-40,000 feet) because of the combination of high dynamic pressures at these altitudes and severe wind shears associated with jet stream occurrences.

Usually all the data available at the present time for application to missile response problems have been obtained by means of balloon-tracking methods. The large-scale wind shear characteristics of the atmosphere which have been obtained from studies of these data [Sissenwine, 1958] have served satisfactorily for initial investigations of the loads and response characteristics of missiles. For studies of the response of complete missile systems and for refinement of design criteria, however, a more detailed knowledge of the wind profile is required. The information obtainable in the balloon-measured profiles is limited by both the accuracy limitations of the tracking equipment and the imperfect

response of the balloon itself as the scale of the wind fluctuations is decreased [Killen, 1960]. It is thus apparent that with the techniques for measuring winds from balloon observations it is not possible to resolve wind variations of the scale that is generally needed for missile systems analysis.

A number of methods of measuring wind profiles have been suggested, including use of angle-of-attack measurements from either an ascending missile [Reisig, 1956] or a dropped object [Weddleton and Larrabee, 1959], the use of multiple balloons and discrete smoke puffs [Chamberlin, Barad, Ely, and Lettau, 1957], and the use of radar tracked chaff-coated balloons. However, these methods for a number of reasons have not provided the small-scale detail needed for missile response studies over the range of altitudes required.

In an effort to provide measurements that include the smaller scale fluctuations of the wind, the National Aeronautics and Space Administration developed a technique based on the use of photographs of visible trails emitted from vertically rising rockets. In addition to providing information on the small-scale wind fluctuations at different altitudes, the smoke-trail technique has the further advantage of providing the measurements along a typical missile trajectory, thus eliminating the differences in the wind structure between the time-space path of a balloon and the near-vertical missile trajectory.

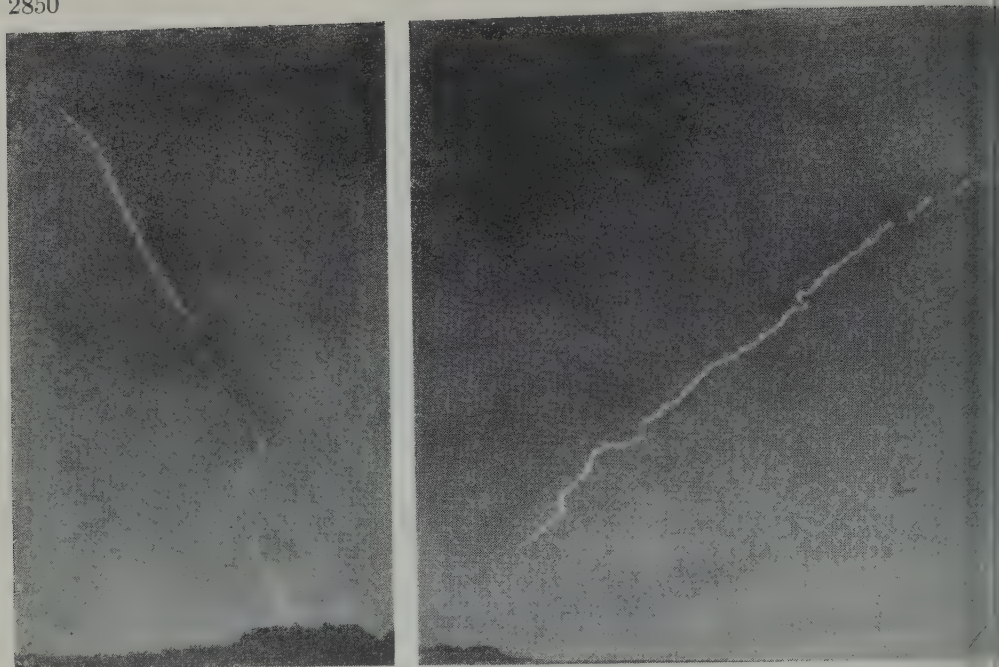


Fig. 1. Photographs from two camera sites of the trail left by the rocket motor.

The purpose of this paper is to describe the smoke-trail technique and to present some wind measurements obtained by this method for comparison with balloon measurements. The method of evaluating the photographs is described, together with a consideration of the accuracies involved.

SYMBOLS

b	length of base line between cameras, ft.
f	principal distance of camera, in.
V_z	component of wind velocity in the Z direction, ft/sec.
x, z	horizontal and vertical film coordinates, in.
x', z'	horizontal and vertical film coordinates for ideal camera, in.
X, Y, Z	right-hand cartesian coordinates with origin at camera I, X axis along the camera base line, and Z axis vertical, ft.
α	azimuth angle at camera from Y axis to point on smoke trail, deg.
Δt	time increment between pictures, sec.
ϵ	dihedral angle at camera base line from horizontal plane to plane containing point on smoke trail, positive upward, deg.

σ_r	root-mean-square error of film coordinates, in.
σ_v	root-mean-square vector error of velocity, ft/sec.

Subscripts:

I	refers to camera I.
II	refers to camera II.

MEASUREMENT TECHNIQUE

Generation of smoke trail. Many motors, particularly some of the solid-propellant motors, produce a visible wake or smoke as a result of the introduction of the combustion products into the atmosphere. An example of such a trail is shown in Figure 1. Motion-picture studies show that the trail reaches equilibrium with the ambient atmosphere within a few seconds, at least in the relatively dense atmosphere below 100,000 feet. This trail, of course, is formed only during the thrusting portion of the missile flight which may not, in all cases, cover the altitude range of interest in the present study. A supplemental method has been developed to produce a missile trail during the coasting portion of the flight for wind shear measurements.

Trails for the present tests were produced by

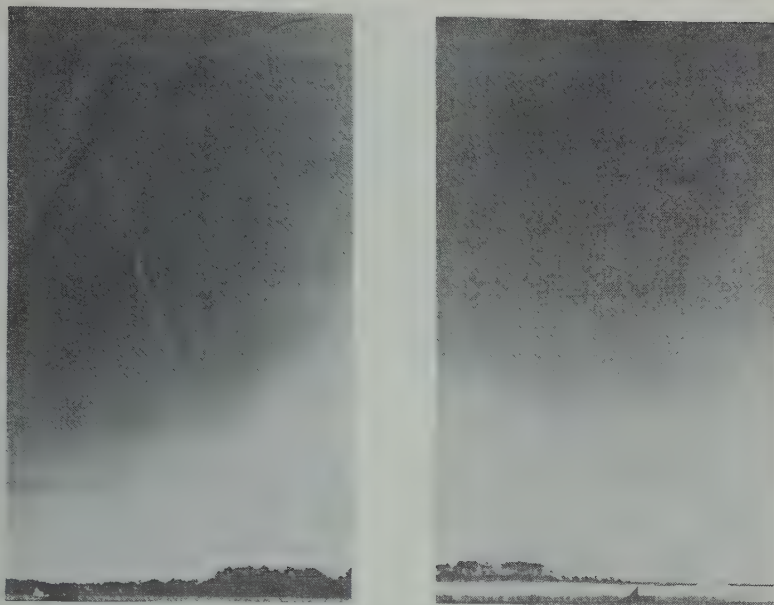


Fig. 2. Photographs from two camera sites of the trail left by a special smoke-generator rocket.

ing a suitable chemical into the atmosphere during the coasting portion of rocket flights. Sulfur trioxide (SO_3) in chlorosulfuric acid (ClSO_3H) was used for this purpose. The solution is known by the chemical warfare name FS. When introduced into the atmosphere as an aerosol, FS combines with atmospheric vapor to form hygroscopic particles of chlorosulfuric acid (HCl) and sulfuric acid (H_2SO_4). The strong affinity of these particles for water causes very rapid condensation of the moisture present in the atmosphere seeded by the aerosol. For the present investigation FS was introduced along the flight path of the rocket and the condensation resulted in a highly reflective trail beginning from the rocket fuel burnout at an altitude of about 5000 feet throughout the coasting portion to apogee near 50,000 feet. An example of the trail produced by this method is shown in Figure 2. This trail was considerably smaller in diameter and more regular in outline than that produced by the thrusting rocket, permitting precise measurements to be made.

The manner in which stagnation pressure is used to discharge the smoke-producing agent from the rocket nose cone used in the present investigation is illustrated in Figure 3. When the vehicle is accelerating, the liquid is in the position shown in Figure 3a. During this acceleration period, air flows into the impact tube,

bubbles through the liquid, and then flows out the discharge tube. After the rocket burns out, aerodynamic drag causes deceleration of the rocket relative to the liquid, and the liquid moves to the position shown in Figure 3b. Stagnation pressure then forces the smoke-producing liquid out the discharge tube. Recombination and condensation processes as described previously form the trail.

The general configuration of the complete vehicle used to produce the artificial trails is shown in Figure 4. The vehicle shown in the figure consists of a Nike booster rocket motor fitted with the special nose cone previously described.

Photography. The principles of the photographic technique are illustrated schematically in Figure 5. Pairs of simultaneous pictures of the trail are taken with cameras I and II at time intervals of a few seconds. Each simultaneous pair of pictures establishes the position of the trail in space; the wind velocities are then determined from the movement of the trail over the time interval between successive pairs of pictures.

As indicated in Figure 5, the cameras are oriented so that the horizontal projections of the optical axes intersect in the vicinity of the launching site. The cameras are elevated so that the field of view includes the horizon and covers

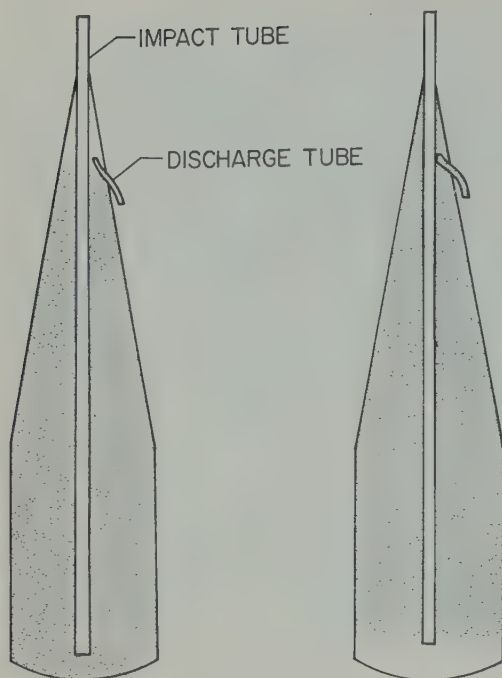


Fig. 3. Nose cone for discharging smoke-producing agent during coasting phase of rocket flight. (a) Rocket accelerating, FS not being discharged. (b) Rocket decelerating, FS being discharged.

altitudes up to approximately 100,000 feet over the launching area. Markers on the horizon within the lower portion of the field of view are used to establish reference marks on the photographic negative.

The cameras currently being used are Air Force type T-11 precision aerial mapping cameras. These cameras provide a 9- by 9-inch negative and are capable of single-frame operation at 3-second time intervals. In order to minimize film distortion, the film is vacuum flattened in the focal plane of the camera. The T-11 camera also provides with each exposure a photograph of a clock, a frame counter, and data panels giving camera serial number, lens serial number, and the calibrated focal length of the lens, plus information entered by the photographer such as location, date, and shutter settings. In addition, lighted and shadowed fiducial marks identify the optical axis of the camera, and calibrated focal length marks along two sides of the picture provide a means of determining the principal distance. Maximum contrast between the white trail and the blue

sky is obtained through use of a Wratten red filter in addition to the permanently installed yellow filter. In the present study, Super Aero graphic film with an ASA index of 11 is used. The cameras are equipped with 7- and 6-inch focal length Metrogen lenses. A shutter speed of 1/150 second has been found to produce satisfactory photographs during daylight in clear sky. Simultaneous photographs at the camera stations are provided by synchronous motor timers which are synchronized by means of a radio signal.

EVALUATION PROCEDURE

As can be seen from the photographs of smoke trails in Figures 1 and 2, a point on the trail in the picture from camera I cannot readily be identified with the same point in the picture from camera II. This identification of points on the trail constitutes the principal special feature of the present method of solving the smoke-trail triangulation problem. This identification might be accomplished in various ways, such as the analog method used in connection with the sodium vapor releases at low altitudes [Manring, Bedinger, and Pettit, 1954]. However, the degree of precision desired for the purposes of the present study is more readily achieved with digital reduction techniques. The following paragraphs describe the method currently being used at the Langley Research Center.

Geometrical relationships and establishment of coordinates of trail. In this section the triangulation method to establish the position of the trail in space from measurements of the coordinates of the film images is described. Figure 6 shows schematically the cameras and the X, Y, Z coordinate system used for position evaluation. It can be seen by inspection of Figure 6 that

$$\tan \alpha_I = X/Y$$

$$\tan (-\alpha_{II}) = (b - X)/Y$$

$$\tan \epsilon = Z/Y$$

and solving these equations yields

$$X = \frac{b \tan \alpha_I}{\tan \alpha_I - \tan \alpha_{II}}$$

$$Y = \frac{b}{\tan \alpha_I - \tan \alpha_{II}}$$

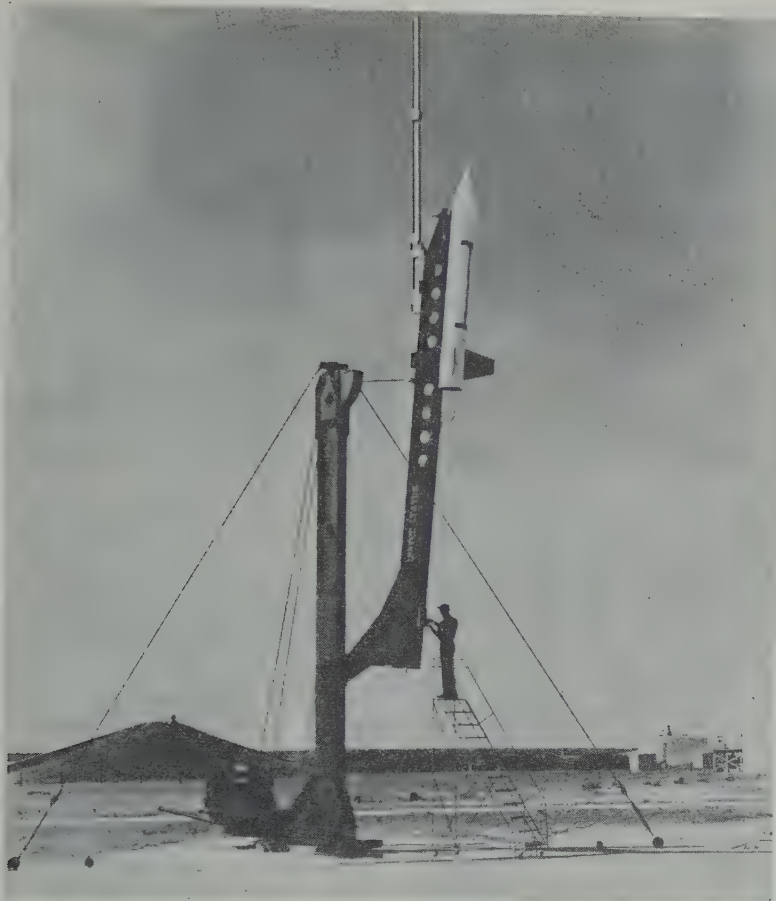


Fig. 4. Nike rocket with smoke-generator nose cone.

$$Z = \frac{b \tan \epsilon}{\tan \alpha_I - \tan \alpha_{II}} \quad (6)$$

These equations are equivalent to the classical maximum baseline solution of the triangulation problem as illustrated, for example, by *Sharp* [1951]. The problem is to apply these equations to obtain the position from the film images.

To illustrate the principle it will be assumed that both axes are pointed horizontally and perpendicular to the base line as illustrated in Fig. 7. Although it is not convenient to align the cameras in this fashion, the same result may be achieved through the use of a coordinate transformation which is described later. Since the object is very far from the camera, the following simple relationships can easily be seen:

$$\tan \alpha = x'/f \quad (7)$$

$$\tan \epsilon = z'/f \quad (8)$$

Since in general there is only one point X, Y, Z on the smoke trail which subtends a given elevation angle ϵ , the same point in the photographs from camera I and camera II is identified by the condition

$$\epsilon_I = \epsilon_{II}; \quad \tan \epsilon_I = \tan \epsilon_{II}; \quad \frac{z_I'}{f_I} = \frac{z_{II}'}{f_{II}} \quad (9)$$

If the focal lengths of the two cameras are equal,

$$z_I' = z_{II}' \quad (10)$$

and a common point will have a common z' coordinate value. For any given value of $z_I' = z_{II}'$ one can measure the corresponding value of x_I' and x_{II}' and from (7) and (8) find $\tan \alpha_I$, $\tan \alpha_{II}$, and $\tan \epsilon = \tan \epsilon_I = \tan \epsilon_{II}$. Substitution of these values into (4), (5), and (6) yields the

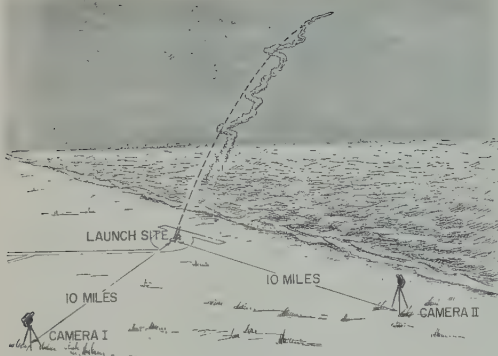


Fig. 5. Camera installation.

X, Y, Z coordinates of the position in space of this particular point on the smoke trail. Following this procedure for a sufficiently large number of z' values will then define the position of the smoke trail with the desired degree of detail.

Film reading. Since the trail is white and the blue sky background appears black through the filters, it is convenient to make readings from the negative, showing a black trail on a clear background. Any loss of accuracy which might occur in the printing and processing of a positive copy is also avoided through use of this procedure. The $x-z$ film coordinates of points along the trail for given pairs of photographic negatives are read with a Wild St.k-1 comparator. To provide rapid coding and recording of the large number of points required, telereader reading heads have been attached to the x and y shafts of the comparator. These reading heads are connected to a Telecorder coder and IBM summary punch providing punchcard recording at the rate of 50 cards per minute. The x and z coordinates of points along the trail are read at intervals sufficiently small so that straight lines connecting the points will completely define the trail position. While the negative is in the same position, the coordinates of the fiducial marks, coordinates of the calibrated focal length marks, and coordinates of the film images of surveyed reference points along the horizon are also read so that camera orientation and film shrinkage can be determined.

Coordinate transformation. As previously mentioned, it is not feasible or desirable in the present application to place the cameras in the orientation which would be convenient for

computation, since to do so and still cover entire trail would require either a great reduction in the base line resulting in a great loss of accuracy or moving the cameras to a much larger distance from the trail with a considerable loss in accuracy plus an intolerable atmospheric visibility requirement. In addition, accurate control of the camera orientation is difficult and expensive. Therefore, the cameras are placed and oriented in such a way as to obtain the best coverage of the trail, and then the film coordinates from the cameras so oriented are transformed to those which would have been obtained from cameras pointed horizontally and perpendicular to the base line.

This transformation consists of the following chronological sequence of steps:

1. Translation in x and z which translates the origin of the coordinate system of the camera to the optical axis of the camera.

2. Scale change in x and z consisting of a division of x and z coordinates by the respective principal distances determined from measurement of the calibrated focal length marks, which corrects for film stretching or shrinkage as well as differences in the calibrated focal lengths of the two cameras.

3. Rotation about the optical axis through the measured swing angle which corrects for misalignments of the swing angle of the camera and/or in the orientation of the negative on the comparator.

4. Rotation downward through the measured

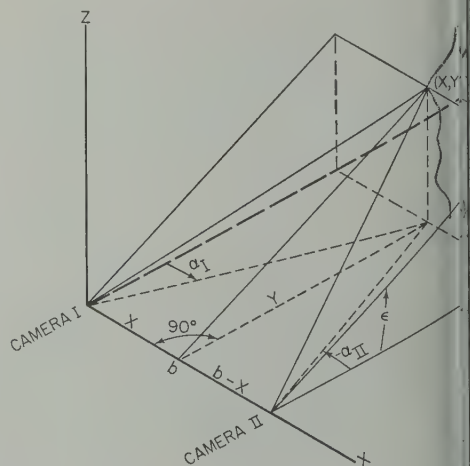
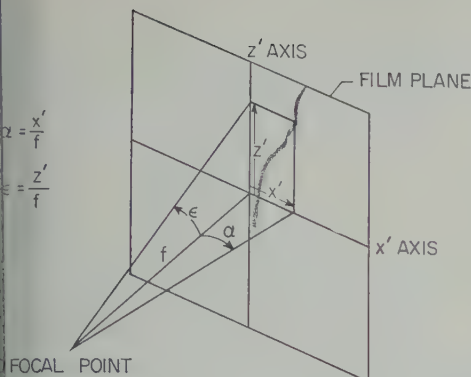


Fig. 6. Basic values used for determining position in space.



7. Camera relationships after coordinate transformation.

tion angle to give the coordinates which have been measured by a horizontal camera.

Rotation about the vertical axis through measured camera azimuth angle to give the coordinates which would have been measured if camera pointing perpendicularly to the baseline.

Merritt [1958] and Brown [1957] for the equations for coordinate transformations of this type.

Calculation of trail position. From the coordinate transformation program the azimuth and elevation angles, α and ϵ , shown in Figure 6 are found for each point read for each of the two cameras. The elevation angle ϵ is the dihedral angle between the horizontal plane and the plane containing the base line between the two cameras and the point (X, Y, Z) on the smoke trail. As previously noted, the elevation angle ϵ thus found is the same for both cameras for any point on the trail. This furnishes the means of identifying the same point in the pictures from two cameras.

The values of $\tan \alpha$ and $\tan \epsilon$ are computed for each point read from each camera by use of equation (8). These values are then entered into the computer as tables of $\tan \alpha$ versus $\tan \epsilon$ and $\tan \epsilon$ versus $\tan \alpha$. Of course it is unlikely that a reading will have been made for the same point from both of the cameras, so it is necessary to obtain some values of the azimuth angles by interpolation. Also, since the bottom of the trail is often indistinct and the visibility and the distance to the trail may differ at the two

camera sites, there may be several data points for one camera corresponding to lower altitudes than any point from the other camera. The computation is begun with the lowest point at which readings have been made from both cameras.

The computer begins with the lowest value of $\tan \epsilon$ from the camera having the greater value for the lowest point read, finds the value of $\tan \alpha$ for this camera, and interpolates a value of $\tan \alpha$ for the other camera. Values of X, Y, Z are then computed from (4), (5), and (6). The computer then proceeds to the next lowest value of $\tan \epsilon$ from either camera, interpolates to find a value of $\tan \alpha$ for the other camera, computes X, Y, Z , proceeds to the next value of $\tan \epsilon$ from either camera, and so on to the top of the trail. The space coordinates are computed for the X, Y, Z coordinate system shown in Figure 6. Thus, the X, Y, Z values corresponding to each point which was read for either of the two cameras are determined. An additional coordinate rotation may be performed to a coordinate system referred to true north or range coordinates.

Calculation of wind velocity. After two successive positions of the smoke trail have been determined, the components of wind velocity are found by simply taking the difference between successive X 's and Y 's at the same Z and dividing by the time interval. Of course, the two successive trail positions will not generally have points at the same Z , and it is once again necessary to interpolate. Since winds are usually desired at equally spaced heights, the interpolations are made at even height values—every 100 feet, for example.

TABLE 1. RMS Vector Wind Velocity Error at $X = Y = Z = 50,000$ Feet as a Function of RMS Film Coordinate Error and Time Increment (Values in table are fps.)

RMS Film Position Error, σ_r , inches	Time Interval between Exposures, Δt , sec					
	1	3	6	30	60	120
0.00004	1.2	0.38	0.19	0.04	0.02	0.01
.0001	2.9	.96	.48	.10	.05	.02
.0004	12	3.8	1.9	.38	.19	.10
.001	29	9.6	4.8	.96	.48	.24
.003	86	29	14	2.9	1.4	.72

ACCURACY

Although with photogrammetry a very high degree of accuracy is possible, careful attention must be given to many sources of error if this potential accuracy is to be achieved. In the particular application under consideration a number of error sources which are usually very important produce little effect because one is concerned only with changes in the position of the smoke trail, rather than with its absolute position at a given time. Thus, even large errors in position will cancel out if they are consistent from frame to frame. Full consideration must still be given to all errors which are random in nature. In the Appendix a number of errors, both systematic and random, are discussed individually, and it is indicated that the systematic errors generally can either be corrected or neglected, as far as wind velocity computations are concerned. The random errors are of such a nature that they may be lumped together into a combined film coordinate error and treated jointly.

Analyses of the errors shows that the rms velocity error is directly proportional to the rms combined film coordinate error, σ_r , is inversely proportional to the principal distance of the camera, and is inversely proportional to the time interval used. It is also a function of the position and height of the trail, the camera orientation angles, and the slope of the smoke

trail. For the installation in use at the Wake Island Range and a nearly-vertical smoke trail, a typical value for the total rms vector error of the wind at 50,000 feet is approximately

$$\sigma_v = 28,800 \frac{\sigma_r}{\Delta t}$$

Particular values of σ_v in fps for some appropriate values of σ_r and Δt are shown in Table 1.

The first value in Table 1, $\sigma_r = 0.00004$ inch (1 micron), is the accuracy of the St.k-1 comparator, and it is a limiting value of accuracy which might be approached by using glass plates with great care in processing and reading. The next value, $\sigma_r = 0.0001$ inch, is the rms error derived from a sample of actual measurements of reference points over the range of temperature and humidity allowed by the air conditioning equipment in the comparator room. Since the reference points have sharply defined images, this value represents the limit of accuracy obtainable with film. The next value, $\sigma_r = 0.0002$ inch, is an estimate of the accuracy obtainable with actual trails under typical conditions. The value of $\sigma_r = 0.001$ inch is the optimum for oscillograph reading equipment used in earlier phases of this study, and $\sigma_r = 0.003$ inch is an average error for this earlier equipment.

Since missiles pass through the wind layers of interest in a fraction of a second, it is desirable to measure the wind over the short

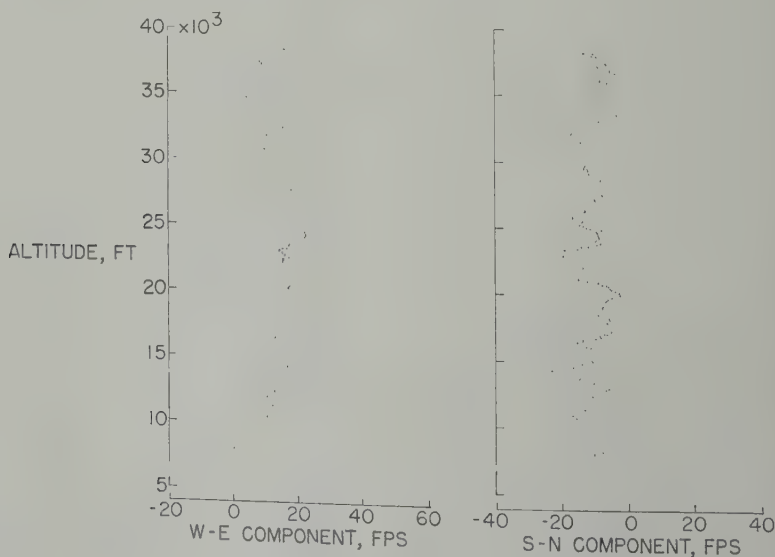


Fig. 8. Smoke-trail winds, 1500 EST, September 21, 1959.

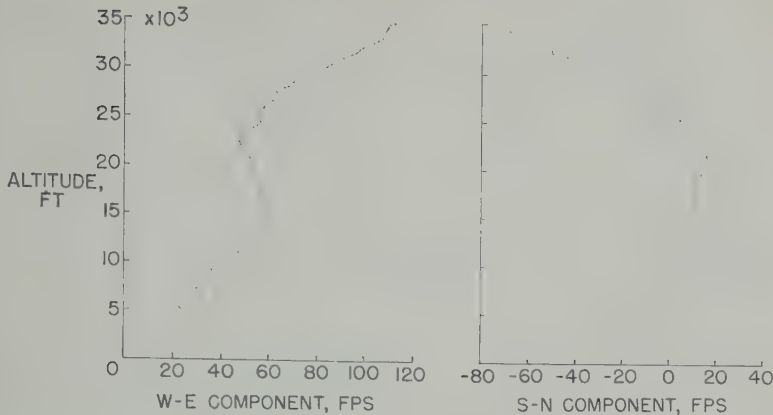


Fig. 9. Smoke-trail winds, 1040 EST, November 4, 1959.

interval consistent with accuracy. For this reason, several time intervals are shown in Table 1. The T-11 cameras are capable of completing a cycle every 3 seconds. The 1-second interval shown can be achieved with a multiple camera installation. The longer time intervals are used to increase the accuracy of the wind velocity with low film position accuracy. In this connection, it may be recalled that the time interval used in radiosonde winds is 2 minutes, the rms accuracy of the winds in the altitude range of interest to the present study is about 0.10 fps [Tolefson, 1956] as compared with values of the order of 0.10 fps in Table 1.

EXAMPLES OF DATA

Figures 8 through 10 illustrate the type of data obtained by the smoke-trail technique from measurements taken at the NASA Wallops Station. In each of these profiles wind velocities have been determined at altitude increments of 1000 feet. This degree of detail in specifying the wind profile is considered adequate for missile response problems encountered to date, but even greater detail might be justified with the more accurate data from smoke trails generated during the coasting phase of a rocket. The values shown in these figures are 1-minute average winds measured with earlier equipment, rather than equipment in current use. The rms error of the values is slightly under 1 fps. As indicated in Table 1, equipment currently in use yields greatly reduced errors for a given time interval. Figure 8 shows the west-east and south-north components of wind velocity at 1500 EST November 21, 1959, evaluated from the trail of

a special smoke-producing rocket at the Wallops Station. For these and subsequent plots, winds blowing from the west and from the south are of the positive sign.

Notice that, even though the winds are light at all levels in Figure 8, disturbances having wavelengths considerably less than 1000 feet can be seen. Also note the random distribution with altitude of the small- and large-scale disturbances. The unsteady aerodynamic loading conditions experienced by a missile in passing through these wind disturbances are of great interest in connection with missile response problems. As a particular example, the combination of the missile flight speed and the spatial distribution of the wind disturbances may result in input frequencies which coincide with some of the predominant structural frequencies. Under such conditions, the system responses lead not only to critical structural loadings but to guidance and control problems as well. Calculations for the loads and responses of missiles for typical flight trajectories through these and other wind profiles are being performed.

Figure 9 shows the profiles of the two wind components determined from the exhaust of a Little Joe rocket launched at 1040 EST November 4, 1959. Here, with moderate winds in the troposphere and a jet stream near the top of the profile, quite appreciable small-scale disturbances are evident throughout the profiles. Figure 10 shows another set of profiles measured from the exhaust of a Little Joe rocket which was launched at 1120 EST December 4, 1959. These profiles show disturbances of both large and small scale throughout the altitude range

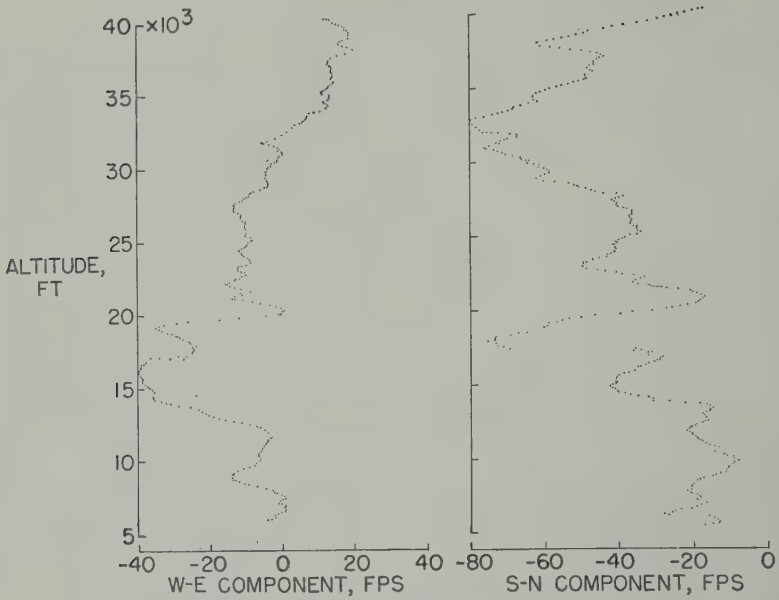


Fig. 10. Smoke-trail winds, 1120 EST, December 4, 1959.

covered. Disturbances of this nature are of particular importance to the missile response problems previously noted.

It is of interest to compare the smoke-trail winds with balloon winds taken near the same time. The smoke trail and balloon winds will not correspond exactly in space and time, since

the smoke-trail winds are measured at a single location and at a single instant of time over the entire altitude range whereas the balloon requires about 40 minutes to reach the top of this range and may drift a considerable distance downstream during its ascent. Also, since the raw sonde balloon winds represent averages over

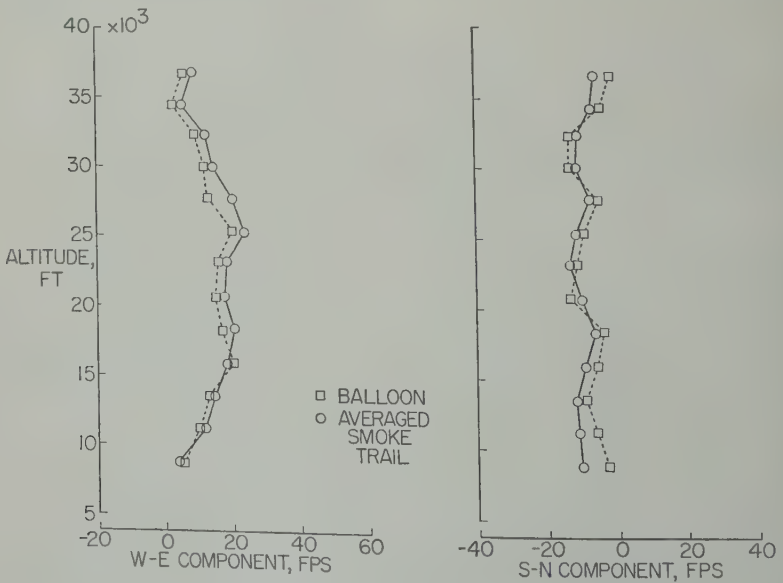


Fig. 11. Balloon and averaged smoke-trail winds compared, 1500 EST, September 21, 1959.

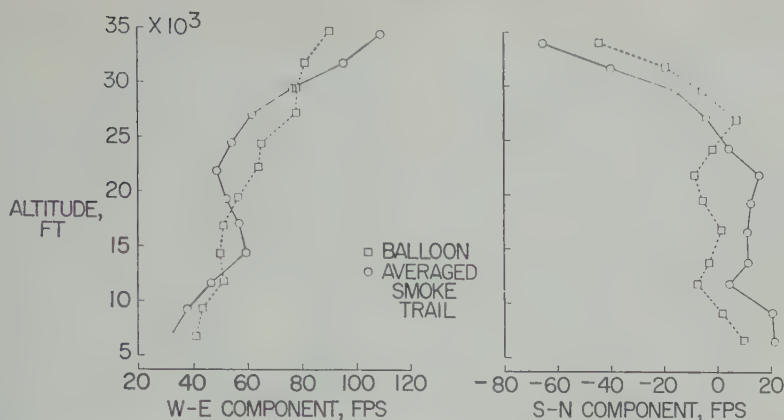


Fig. 12. Balloon and averaged smoke-trail winds compared, 1040 EST, November 4, 1959.

ade interval of about 2000 feet, the profile d be quite different in appearance from the ed smoke-trail profiles.

er comparison with radiosonde balloon winds, rofiles shown in Figures 8 through 10 have averaged over the same altitude intervals e radiosonde winds. These average winds own in Figures 11 through 13 along with rresponding radiosonde winds. In the mber 21, 1959, example of Figure 11, the at was fired when the radiosonde balloon at 30,000 feet, and, because light winds oiled, the balloon remained near the launch- te. In the other two examples (Figs. 12

and 13) the time of rocket firing was controlled by other considerations; however, the balloon winds were measured within 6 hours of the rocket launching time.

In the case of Figure 11 the average winds measured by the two techniques are quite similar; in the other cases (Figs. 12 and 13) the differences are no larger than might be expected for the time and space separations involved. Thus, it can be seen that, insofar as the large-scale, average values are concerned, there is no conflict between the smoke-trail and the balloon winds. Differences are due primarily to the greater detail of the smoke-trail wind profile.

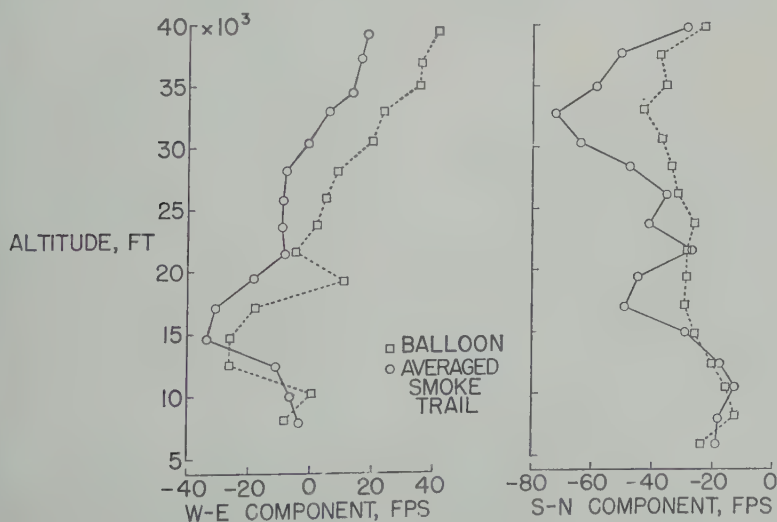


Fig. 13. Balloon and averaged smoke-trail winds compared, 1120 EST, December 4, 1959.

CONCLUDING REMARKS

A method has been presented for determining a detailed profile of the horizontal wind velocity over a large range of altitudes. Data of this type are urgently needed for the determination of dynamic loads on ascending missiles. No other currently operational method is known which is capable of providing sufficient detail and accuracy over the range of altitudes needed for missile response studies. Examples of profiles measured by this method demonstrate the presence of disturbances having wavelengths which will excite the fundamental modes in existing and proposed large, flexible missiles. In addition to missile applications, such detailed profiles also contribute to basic understanding of the structure of the atmosphere.

APPENDIX

Error Analysis

In the analysis of the accuracy of the smoke-trail technique of wind determination, a number of sources of error must be considered. In the following paragraphs various error sources are considered individually. It will be seen that the systematic errors, which tend to cancel out in the subtraction of successive trail positions, are small compared with the random errors, which do not generally cancel out. It will also be seen that the nature of the random errors is such that they can be combined into a single total random film coordinate error, σ_r . This makes possible the development of error equations in terms of a combined error.

Systematic Errors

Surveying errors. Errors in the surveyed positions of the cameras and control points are frequently critical in photogrammetric methods. In the present case, however, relatively large surveying errors can be tolerated since only the movement of the trail, not its absolute position, is of interest. Those control point errors which produce the same position error in successive pictures will cancel out in the computations of wind velocity. An error in the bearing of the base line between the two cameras will result in the same angular error in the wind direction, and an error in the length of the base line will produce the same percentage error in the wind speed. Thus, for example, an error of 100 feet

in the 100,000-foot base line used in the present study would result in an error of only 0.1 per cent in the wind speed.

Errors in the alignment or position of the cameras. Errors in the azimuth, elevation, swing angles of the cameras are all corrected in the coordinate transformation by computing each of these angles from the positions of images of known reference points along the horizon. Thus, the only errors in alignment remaining are those due to reading the coordinates of the reference points or to errors in the surveyed position of the reference points.

Lens distortion. The manufacturer's specification for the distortion of the Metrogon lens in the T-11 camera varies between about $+0.1$ mm and -0.1 mm [*Am. Soc. Photogrammetry*, 1954], producing maximum position errors of about ± 100 feet for the present camera setup. However, the error in wind velocity depends on the difference of error between successive pictures. In the above reference it is shown that the greatest gradient of error due to lens distortion is about 0.1-mm change of error in 20 mm film distance, corresponding to a wind velocity error of 0.5 per cent. Similarly, the height error for any wind measurement might be ± 100 feet, but the error in the increment of height between any two measured winds would not exceed 0.5 per cent.

Focal plane distortion. In the T-11 camera the film is held by vacuum against a slow-precision-finished focal plane plate. Camera specifications require that the surface of the plate be within 0.001 inch of a true plane over the entire plate. The same plate is used for all exposures in the same camera, however, and the deviations from the true plane do not change abruptly, the errors in wind velocity from this source will be quite small. The sensitive emulsion is on the front of the film, and there may be small deviations due to inequalities in the thickness of the film base. Since these variations in thickness are of a random nature and are the same for successive pictures, they will be included with film distortion in the treatment of random errors.

Film shrinking or stretching. Owing to changes in temperature and humidity during storage, processing, and handling, shrinkage or expansion of the film base may occur. Since the film is placed under tension at various times during

camera film transport, developing, drying, handling, these dimensional changes may be the same in the direction along the roll film as across it. These effects are corrected by dividing the film coordinate x values by the principal distance derived from measurements of the calibrated focal length marks in the x direction. Similarly, the z values are divided by the principal distance derived from measurement of the calibrated focal length marks in the z direction.

Other camera errors. A number of possible error sources exist in imperfections in the camera, such as dimensional changes resulting from rough handling or from temperature changes after the camera is calibrated; however, these are all small, and change slowly with changes in the position of the film image and are consistent from frame to frame. It is reasonable to expect that all these effects will produce a smaller error in wind velocity than, for example, the lens distortion, and they are therefore neglected.

Atmospheric refraction. Atmospheric refraction causes light rays to curve in the vertical plane. Correction for this curvature presents a difficult problem in many photogrammetric applications. However, the effect of mean value of the vertical gradients of density and humidity over the relatively small distances involved in the present study is negligible. The variations of refractive index due to surface heating and turbulence are included in the random error.

Earth curvature. The calculations presented in this paper are performed in tangent plane coordinate systems. Over the distances involved, the differences between this system and an earth-centered spherical system are small. For application to greater distances, an additional coordinate transformation to an earth-centered spherical system can be performed, if desired.

Random Errors

Reading errors. The nominal accuracy of the St. k-1 is ± 1 micron, or 0.00004 inch. This error is considerably smaller than the other random errors. An additional error associated with reading is the error of the operator in determining the center of the trail. This reading error is small for the special smoke rockets with a diameter of about 10 feet, but it may be appreciable for exhaust trails from large

rockets during the burning phase, where the diameter of the trail may exceed 100 feet.

Film distortion. Distortion of the film image may occur as a result of nonuniform stretching or shrinkage of the film base, variations in the thickness of the film, imperfect contact with the camera focal plane surface on the mounting plate of the comparator, or creeping of the emulsion relative to the film base. The errors resulting from this distortion will affect the over-all accuracy in the same manner as reading errors. The magnitude of the distortion depends on the type of film used, on the temperature and humidity during photography, storage and handling, and on the developing and drying technique, and it may vary considerably from case to case.

Resolution limits. Poor resolution or lack of contrast with the sky background may increase the reading error above the nominal value of the comparator, particularly under unfavorable weather conditions or sun positions.

Computation errors. At all stages of the computations a sufficient number of digits is used to keep the round-off error small compared with the reading error.

Timing errors. Since the wind components are found by dividing the displacement of the trail by the time interval, an error in the time interval between pictures will produce an equal percentage error in the wind velocity. With the synchronous timers used in the present study, this timing error is only a fraction of a per cent for even the shortest time intervals and is neglected.

Tracer errors. The 'smoke' particles in the trail made by the FS smoke generator are small water droplets similar to those in stratus clouds or airplane condensation trails. In the natural exhaust trail, solid particles of various combustion products may exist in addition to water droplets. All these particles have negligible fall velocities and near-perfect response to all scales of motion of orders larger than the particle sizes. For the rocket exhaust trails it is necessary to wait a few seconds for the trail to lose its initial momentum; however, motion pictures of the exhaust trails of rockets show that this initial momentum is lost almost immediately, at least in the relatively dense air below 100,000 feet.

Refraction index fluctuations. Surface heating and turbulence near the earth's surface produce

movement of the images of distant objects near the surface. This is not appreciable even during summer daytime conditions except for elevation angles within a few degrees of the horizon, but it will affect the smoke trail in the lowest thousand feet or so and the images of reference points on the horizon. These effects can be reduced by choice of the time of day and by increasing exposure time. For the reference points, the effects can be reduced by averaging readings from several points or from the same points at several times.

Effect of vertical air motion. In the computation of the wind velocities it is assumed that the horizontal speed of the particles composing the trail is given by the change in position of the trail at a given height. This will be exactly true if the air, and consequently the trail, has no vertical motion, and it will also be true if the trail is vertical. However, if the trail has appreciable slope, vertical motion will cause a horizontal displacement of the portion of the trail which passes through a given height. The magnitude of this displacement is

$$V_z \Delta t \sqrt{(dX/dZ)^2 + (dY/dZ)^2}$$

and the resulting error in the wind is

$$V_z \sqrt{(dX/dZ)^2 + (dY/dZ)^2}$$

Inasmuch as it is not feasible to determine the vertical velocity at all points along the smoke trail, it is difficult to assign a numerical value to this error. However, the value of vertical velocity in clear air below 100,000 feet is generally quite small (less than 1 fps), except in turbulent air where it may be much larger. (Turbulence at the level of the smoke trail would, of course, produce a rapid break-up of the smoke trail.) It is also possible, in some cases, to find an indication of the vertical velocity from the movement of sharp bends in the smoke trail. This procedure did not indicate appreciable vertical velocity in any of the examples shown in the present paper.

REFERENCES

- American Society of Photogrammetry, *Manual of Photogrammetry*, 2nd ed., George Banta Publishing Co., Menasha, Wis., 1952.
- Brown, D. C., A treatment of analytical photogrammetry with emphasis on ballistic camera applications, *Radio. Corp. Am. Data Reduction Tech. Rept. 39, AFMTC TR 57-22, ASTIA 144*, August 20, 1957.
- Chamberlin, L. G., M. L. Barad, R. Ely, and F. Lettau, Loeser technique of wind profile determination—Daytime smoke-puff and night Pibal swarm measurements, in *Exploring the Atmosphere's First Mile*, vol. 1, edited by H. Lettau and Ben Davidson, Pergamon Press, New York, pp. 276-292, 1957.
- Killen, G. L., Balloon behavior experiments, *Rept. 2093, Army Signal Research and Development Laboratory*, January 15, 1960.
- Manring, Edward, J. F. Bedinger, and H. B. Phipps, Some wind measurements in the upper atmosphere using artificially generated sodium clouds, *Geophys. Research*, 64(6), 587-591, 1959.
- Merritt, E. L., *Analytical Photogrammetry*, Pitman Publishing Corp., New York, 1958.
- Reisig, G. H. R., Instantaneous and continuous measurements up to the higher stratosphere, *Meteorol.*, 13(5), 448-455, 1956.
- Sharp, H. O., *Practical Photogrammetry*, The Macmillan Co., New York, 1951.
- Sissenwine, Norman, Windspeed profile, windshear, and gusts for design of guidance systems for artificially rising air vehicles, *Air Force Survey Geophysics*, no. 57, AFRCR-TN-54-22, December 1954.
- Sissenwine, Norman, Development of missile climb wind profiles for Patrick Air Force Base, *Force Surveys in Geophysics*, no. 96, AFRCR-TN-58-216, ASTIA 146 870, March 1958.
- Tolefson, H. B., An investigation of vertical shear intensities from balloon soundings for application to airplane and missile-response problems, *NACA TN 3732*, July 1956.
- Weddleton, R. L., and E. C. Larrabee, Measurement of winds shear, *Fifth Quarterly Rept.*, Air Corps contract DA-36-039 SC-75064, September to November 30, 1959, United Research Laboratories, Cambridge, Mass., 1959.
- (Manuscript received May 1, 1961; presented at the Joint Meeting of the American Meteorological Society and the American Geophysical Union, Washington, D. C., April 19, 1961.)

Mean Roughness Factor as a Function of Wind Velocity

ALLEN H. SCHOOLEY

*U. S. Naval Research Laboratory
Washington 25, D. C.*

Abstract. Roughness factor of the facets of a wind-disturbed water surface is defined as the ratio of flatness tolerance to facet length. The mean roughness factor for all facets in the upwind-downwind direction is about 0.095 at 10 knots wind velocity and about 0.135 at 20 knots when measured by means of a short fetch water-wind tunnel. Using data presented in this paper, approximate calculations have been made on the nature of the function relating radar beam depression angle with the upwind-to-downwind ratio of radar reflection back from the rough sea surface. The results are in qualitative agreement with the meager existing experimental data. It is predicted that there is an interesting minimum in the upwind-to-downwind ratio in the region of 70° depression angle.

Background. The roughness of the facets of wind-disturbed water surface has been expressed in terms of a dimensionless roughness factor [Schooley, 1961]. Roughness factor was defined as the ratio of flatness tolerance to facet length. These last two terms are defined in Figure 1. The heavy irregular line represents a wave profile with the wind blowing from left to right. The light solid line at an angle θ with respect to the mean surface level, is tangent to one point on the wave profile, represents the slope at this point. Very short electromagnetic waves striking the surface at angles to the tangent will be reflected back from a small facet of the water surface centered on this point. The size of the facet will be determined by its flatness, measured in terms of a fraction of a wavelength. Flatness tolerance has been defined as $1/10$ of a wavelength. In Figure 1 flatness tolerances (a) and (b) are indicated, together with the corresponding facet lengths. Figure 2 shows the profile of a water wave produced by a 16-knot wind blowing from left to

right. The channel is about 10 cm wide and 70 cm long. In the illustration the boundary between the air, the water, and the front transparent wall of the channel was retouched to make it more visible. Full-size pictures similar to Figure 2 were used in taking the experimental measurements which have been reported [Schooley, 1961]. The slope and facet lengths for chosen flatness tolerances were measured graphically at $1/2$ -cm intervals along many wave profiles for various wind velocities. The results were averaged and combined, resulting in Figure 3. Here the ordinate is the average roughness factor for flatness tolerances of 1 mm, 3 mm, 5 mm, and 10 mm, and the abscissa is the surface slope in degrees. Curves for wind velocities of 10, 15, and 20 knots are shown.

It is interesting to note in Figure 3 that in the general region between $+10^\circ$ and -10° surface slope, the average roughness factor is greatest for the high wind velocities. For slopes outside this general region, the situation is reversed. This leads to the question whether the mean roughness factor for all slopes is an increasing or a decreasing function of wind velocity. It seems logical that mean roughness factor should be defined so that it is an increasing function with increasing wind velocity. This fact is not immediately evident in Figure 3 because the relative probability of the various slopes has not been considered. It is known that the probability of a facet of the surface having a slope in the region around zero is greater than

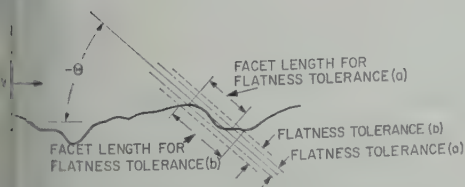


FIG. 1. Flatness tolerance determines facet length.

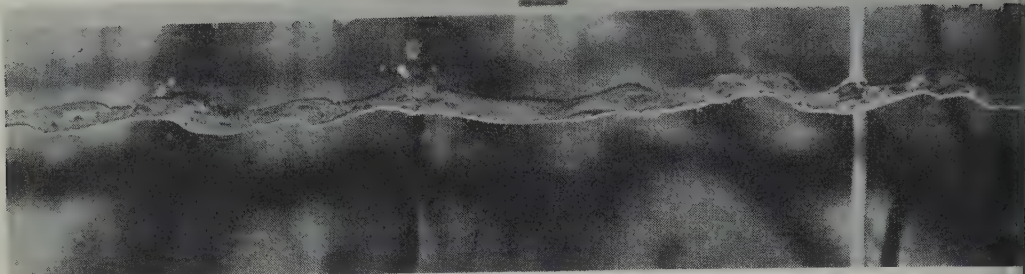


Fig. 2. Profile of a wave in transparent water-wind tunnel.

for very high slopes where the probability all but vanishes [Cox and Munk, 1954; Schooley, 1954].

Slope probability and mean roughness factor. Figure 4 gives previously unpublished data relating the cumulative probability of various surface slopes for various wind velocities. The experimental points were taken from a multiplicity of cumulative probability curves similar to that of Schooley [1958]. By means of Figure 4 it is possible to construct Figures 5, 6, and 7, which relate the probability of various slopes for wind velocities of 10, 15, and 20 knots, respectively. These last three figures represent an upwind-downwind slice of the distribution function. The distribution function is really three dimensional because facets of the water surface tilt in the cross-wind direction as well as in the upwind-downwind direction. For this paper, however, only the upwind-downwind slice of the distribution will be considered because only upwind-downwind experimental data have been taken.

Let it be assumed that a slice of small constant width is considered, having the distribution of slopes given in Figure 6. If we let p_{5i} represent the probability density of slope at the midpoint of any 5° interval $5i \leq p \leq 5(i + 1)$, and R_{5i} represent the value of the average roughness factor at the midpoint of the same 5° interval, then the mean roughness factor for all slopes, \bar{R} , will be given by

$$\bar{R} = \sum_{i=10}^{i=9} p_{5i} R_{5i}$$

where, because of normalization,

$$\sum_{i=10}^{i=9} p_{5i} = 1$$

and where the limits cover the range of experimental data which is between -50° to $+50^\circ$. This calculation gives the center point on the curve in Figure 8. The other points were calculated in a similar way using Figures 5 and 7 together with the proper curves from Figure 4.

Figure 8 shows that the mean roughness factor of a wind-disturbed water surface is in fact an ascending function of wind velocity. Below 10 knots the curve must come down abruptly and lose its significance because a threshold wind velocity is approached where waves are no longer generated.

Although Figure 8 shows the mean roughness factor to be an ascending function of wind velocity, it is the basic data from Figure 3 that are useful in considering some practical problems. Hence, there is some doubt that the name 'roughness factor' is a good choice for describing the ratio of flatness tolerance to facet length. However, no alternative name is suggested.

Radar sea clutter results. The work on relating surface flatness tolerance to facet size has started several years ago in an attempt to be-

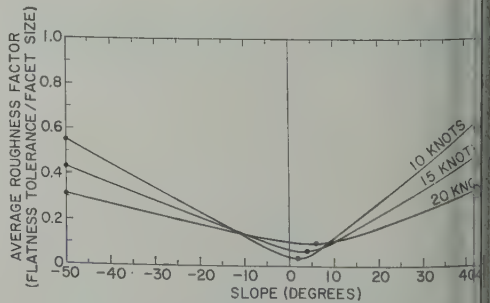


Fig. 3. Average roughness factor versus slope for various wind velocities.

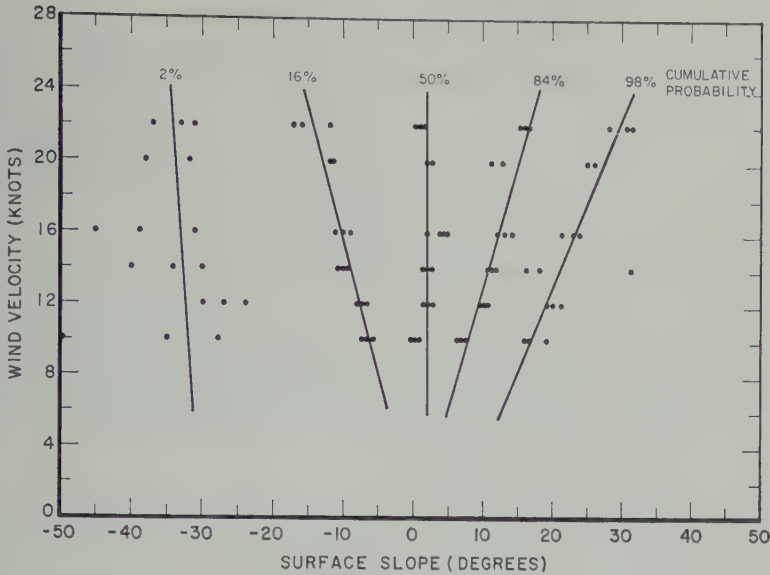


Fig. 4. Cumulative probability versus facet slope and wind velocity.

understand the phenomenon of radar 'sea clutter.' Sea clutter is the irregular signal return that a radar gets from the rough reflecting surface of the sea [Katzin, 1957]. It is interesting to note that some aspects of radar sea clutter may be calculated, for the first time, using the data of Figure 3. Using, in addition, the information obtained in the 15-knot curve of Figure 3, a rather lengthy approximate step-by-step calculation was made which gave the results shown

in Figure 9. The approximate theory used in calculating the results shown in Figure 9 is being refined and is the subject of a paper now in preparation. The new paper will also show the effects of varying wind velocity and flatness tolerance on the upwind-to-downwind ratio. Figure 9 shows the calculated upwind-to-downwind ratio of radar sea clutter versus the beam depression angle of an airborne radar with respect to the horizontal. In this case, the wind velocity is 15 knots and the wavelength 3 cm, corresponding by definition to a flatness tolerance

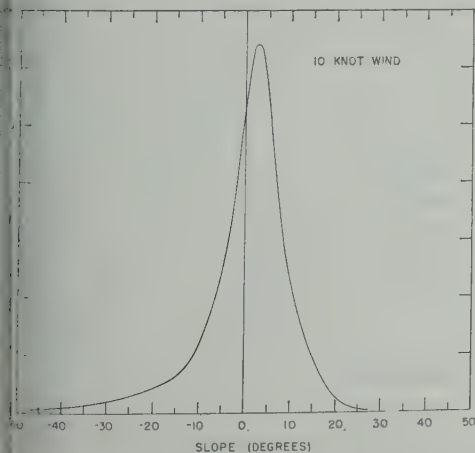


Fig. 5. Slope probability for a 10-knot wind.

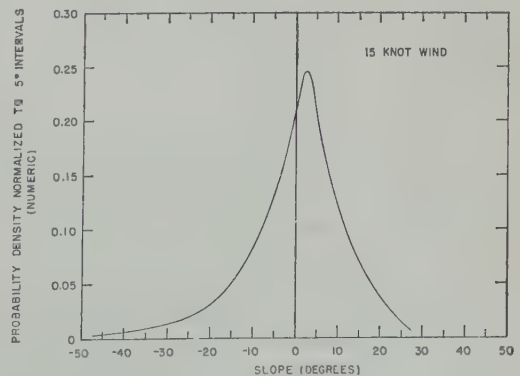


Fig. 6. Slope probability for a 15-knot wind.

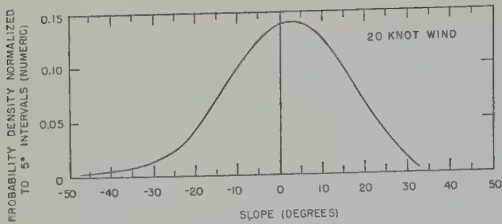


Fig. 7. Slope probability for a 20-knot wind.

of 3 mm. This calculated curve checks observations that sea clutter when 'looking' upwind is higher than when looking downwind for a 3-cm-wavelength radar when the depression angle is small [Macdonald, 1956]. The curve also goes to zero at 90°, which corresponds to looking straight down. Under this circumstance upwind and downwind lose their meanings and the curve should go to zero on the decibel scale. The region around 70° is most interesting. Here, the sea clutter is greater when looking downwind than when looking upwind. This has not been shown before, and there are no known accurate radar measurements at large depression angles to check this prediction experimentally. However, the effect appears reasonable and is probably due to large specular reflections off the broad gentle slopes on the left part of the waves that are shown in Figure 2. Optical measurements made by the author have confirmed the negative upwind-to-downwind ratio in the region around 70°.

Figure 9 also should be applicable to Sonar surface reverberation with the following modi-

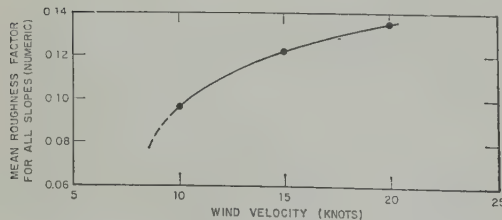


Fig. 8. Mean roughness factor as a function of wind velocity.

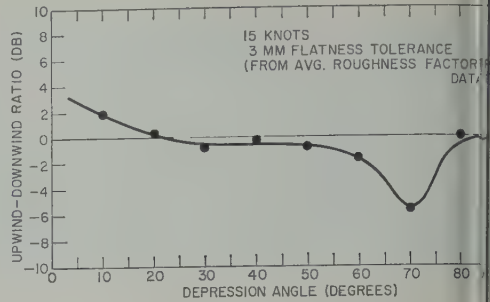


Fig. 9. Calculated upwind-to-downwind ratio vs beam depression angle for a 3-mm flatness tolerance, corresponding to a radar wavelength of 3 cm, and a 15-knot wind.

fications. Since the Sonar beam looks up to the surface rather than down on the surface, the abscissa should be relabeled 'Elevation Angle (Degrees).' Also, the curve should be a mirror image of itself around the zero ordinate. Because of the underwater geometry, the upwind-to-downwind ratio for the Sonar surface reverberation should be negative at the low elevation angles and should become positive in the region around 70°.

REFERENCES

Cox, C. S., and W. H. Munk, Statistics of the sea surface derived from sun glitter, *J. Marine Research, Sears Foundation*, 13, 98-227, 1954.
Katzin, Martin, On the mechanisms of radar sea clutter, *Proc. IRE*, 45, 44-54, 1957.
Macdonald, F. C., The correlation of radar sea clutter on vertical and horizontal polarization with wave height and slope, *IRE National Convention Record*, part 1, 1956.
Schooley, A. H., A simple optical method for determining the statistical distribution of water surface slopes, *J. Opt. Soc. Am.*, 44, 37-40, 1954.
Schooley, A. H., Probability distribution of water wave slopes under conditions of short fetch, *Trans. Am. Geophys. Union*, 39, 405-408, 1958.
Schooley, A. H., Relationship between surface facet average facet size, and flatness tolerance of wind-disturbed water surface, *J. Geophysical Research*, 66, 157-162, 1961.

(Manuscript received May 6, 1961; presented at the Forty-second Annual Meeting, American Geophysical Union, Washington, D. C., April 19, 1961.)

How Dry Is the Sky?

MURRAY GUTNICK

*Geophysics Research Directorate
Air Force Cambridge Research Laboratories
Bedford, Massachusetts*

Abstract. Measurements over Western Europe (by British investigators) indicate a uniformly dry lower stratosphere, with frost points around -82°C and mixing ratios of about 2 milligrams/kilogram, while other observations over North America imply frost points from -60° to -85°C , with mixing ratios varying up to several hundred milligrams/kilogram. Since no theory is able to explain all the observations, they must be re-examined for possible bias or error.

Introduction. Dry-sky advocates argue tediously, in person and in print, with others who think the heavens are more humid. At issue is whether the mixing ratio in the stratosphere is only a few milligrams per kilogram, or 10 or 100 times greater. In this paper an attempt is made to summarize some of the arguments and evidence.

Despite the apparently insignificant amounts of water vapor involved, this 'tempest in a teapot' has both practical and theoretical implications. Over a 200-km path at 9 km, an order of magnitude difference in mixing ratio, even a minute amount of water vapor, may determine whether an infrared detector can pick up the radiation emitted by an intercontinental missile during the boost stage. Also, the stratospheric circulation theories based on certain observed stratospheric frost points would be untenable if the observed frost points proved to be in error.

Of all the interrelated ways of expressing stratospheric water vapor content, the mixing ratio seems best suited for discussion of stratospheric moisture. Since it is a ratio of masses, it is dimensionless. The usual unit of expression, milligrams of water per kilogram of other atmospheric constituents, is equivalent to parts per thousand. In the stratosphere, however, water vapor exists in only a few milligrams per kilogram, or in parts per million. For simplicity of expression and notation, all mixing ratios in this paper are expressed in mg/kg, which may be interpreted as 10^{-3} parts per million.

British ascents. The foremost investigators and data collectors of stratospheric humidity in

the recent past have been the British. Although some isolated American measurements [Barrett, Herndon, and Carter, 1950] were made in 1949 and 1950, the British began taking systematic moisture measurements [Helliwell, Mackenzie, and Kerley, 1956, 1957; Murgatroyd, Goldsmith, and Hollings, 1954, 1955; Tucker, 1957] in the lower stratosphere in the early 1950's. Using aircraft platforms and manually operated frost-point hygrometers, they flew at maximum altitudes of 14 or 15 km. The first ascents were over southern England; later ones were over the Sudan [Goldsmith, 1954], Tripoli [Helliwell and Mackenzie, 1956], and Norway [Brewer, 1955]. In 1956-1957 the British made some 'round-robin' flights [Helliwell, 1960] ranging in latitude from 40° to 67°N .

Almost all ascents in the lower stratosphere found the mixing ratio above about 150 mb to be 2 mg/kg (frost points between -80° and -85°C). These observations at 14 and 15 km were remarkably constant regardless of time of year or geographical location.

Aircraft ascents [Kerley, 1961] over Nairobi (1°S), Aden (12°N), and Bahrain Island (25°N), in June 1958, suggest that the mean summer frost points over Nairobi and Aden are approximately the same between 6 and 12 km, and about 12° to 15°C warmer than over England. Above 12 km, the difference decreases with height so that at 15 km the frost points are about the same (-80°C) as over England.

The British theorized [Gutnick, 1960] that in the lower stratosphere the mixing ratio remained constant with altitude at about 2 mg/kg; this, of course, implies that the frost point must

decrease with height. To account for the extreme dryness of the lower stratosphere, *Dobson* [1956] and *Brewer* [1949] postulated a model of the stratospheric moisture mechanism in which air passes regularly through the tropical tropopause via a meridional circulation. That is, the air ascends in the tropics, is cooled, and is dried out by condensation. This air penetrates the tropical tropopause (at about 100 mb, frequently with temperatures of -83°C) and flows poleward.

If saturated at the tropical tropopause, stratospheric air moving poleward and descending slightly would arrive over England with frost points consistent with those measured with the British frost-point hygrometer. To complete the circulation, the poleward moving stratospheric air must descend into the troposphere at higher latitudes.

The extreme dryness of the lower stratosphere over England has been confirmed on at least two occasions by spectrographic methods [*Houghton and Seeley*, 1960; *Yarnell and Goody*, 1952]. However, *Barclay, Elliott, Goldsmith, and Jelley* [1960], using a nitrogen-cooled vapor trap, found a mixing ratio of 37 mg/kg at 27 km over England; this value is some 19 times as large as had been theorized hitherto by dry-sky advocates.

American ascents. Most American ascents for determining stratospheric moisture have been made by balloon-borne automatic frost-point hygrometers which on occasion have exceeded 30 km in altitude. In contrast to the British ascents, they indicate a wide range of humidities [*Sissennwine and Gutnick*, 1960] both at the same altitude as the British ascents and at higher altitudes in the stratosphere; Figure 1 gives a typical sample of American stratospheric humidity ascents. For example, J. Brown of Ballistic Research Laboratories, Aberdeen Proving Ground, Maryland, made a number of simultaneous ascents at latitudes 10° to 80°N at 1200 GCT on April 8, 1960 (Dewpoint Cross Section-Albrook AFB CZ to Alert, Ellsmere Island 1200 UT, unpublished data). At the 150-mb level (approximately 14 km) and between 40° and 75°N , the frost points ranged from -67° to -85°C and the corresponding mixing ratios from 1 to 17 mg/kg.

On the other hand, measurements of the solar infrared spectrum over New Mexico from about 13 km by *Gates, Murcray, Shaw, and Herbold*

[1958] agree with the British dry-sky theory and spectroscopic analysis of the layer from 10 to 28 km over New Mexico by *Murcray, Murcray, Williams, and Leslie* [1960] yielded essentially the same result.

Near Fairbanks, Alaska, in October 1958 spectrographic measurements from a U-2 aircraft of the Air Force Cambridge Research Laboratories (J. Strong and J. N. Howard, unpublished manuscript, 1960) indicated that the precipitable water above 11 km (240 mb) was 23 microns; above 20 km (60 mb) it was 13 microns. These values correspond to mixing ratios of 9 mg/kg for the entire atmosphere above 11 km and 26 mg/kg for that portion above 20 km.

To illustrate the wide divergence in moisture content reports, mixing ratios (over ice) and frost points at 14 to 15 km for typical British and American ascents are given in Table 1. For those American results for which the exposure at 14 km was not available, the IC Standard Atmosphere pressure of 142 mb was used. Hence, some of the mixing ratios in Table 1 are only approximations and are marked by asterisks.

Discussion. Although the American soundings showed wide variations in moisture content at any given stratospheric level, most (but not all) had one factor in common: mixing ratio increased with altitude [*Gutnick*, 1960] from about 16 to 32 km (the highest altitude to which analysis could be extended). In one such ascent [*Mastenbrook and Dinger*, 1961] the mixing ratio at 30 km actually exceeded that at 12 km by more than a factor of 4 (Fig. 1, curve W).

Some investigators suggest that the apparent variation can be reduced by selecting only the data which are considered reliable. However, no substantial agreement has been reached as to which data are reliable and which are not. One group considers reliable only those frost-point data that show agreement between the ascent of the hygrometer and its subsequent descent. Others point out that there are no criteria for good and poor agreement; hence the selection would be purely subjective. More than 3 hours usually are required for ascent and descent of a run that reaches 30 km; during this interval the balloon may have drifted about 200 km from the launching point, causing only a relatively large temporal difference

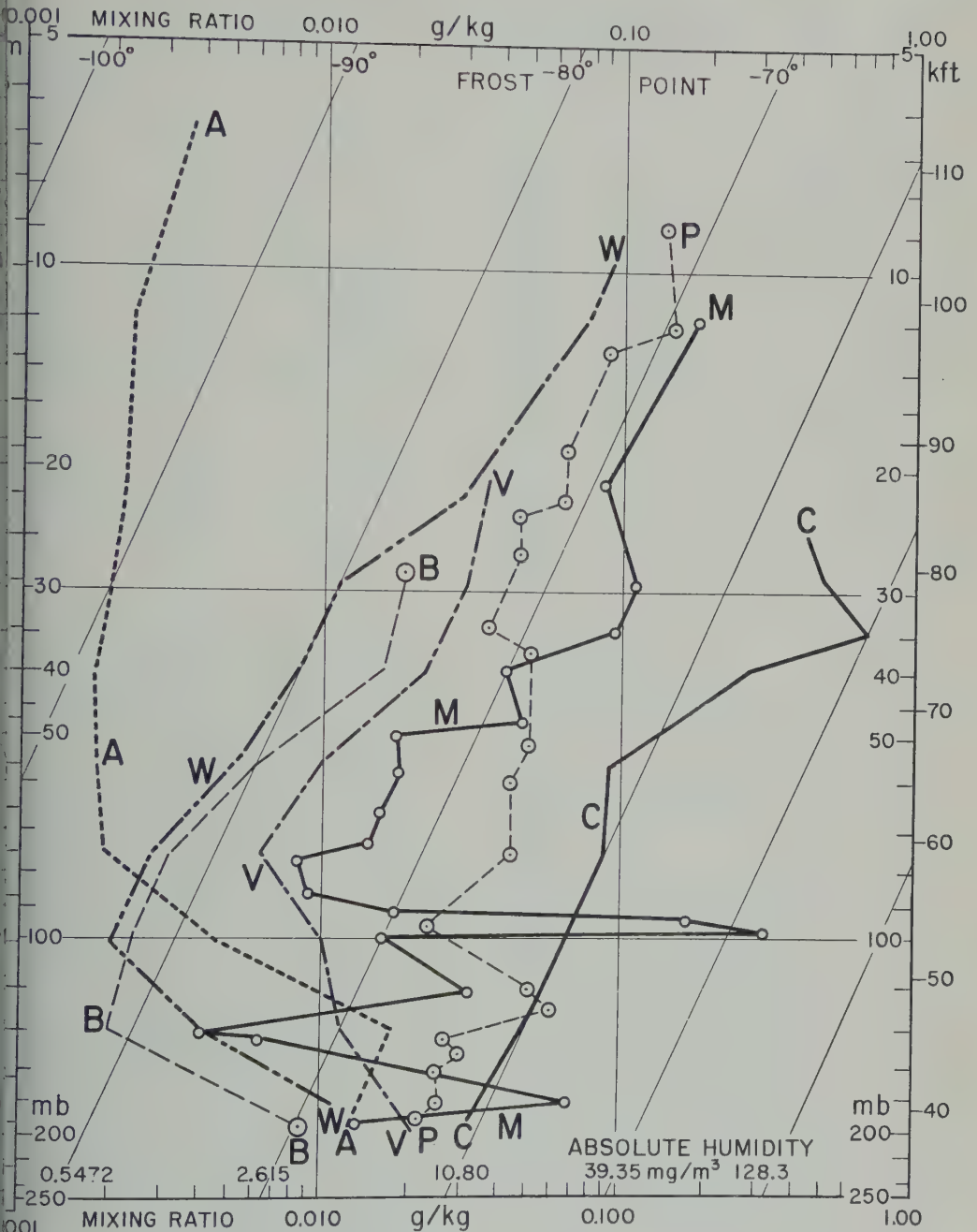


Fig. 1. Typical stratospheric humidity soundings over North America, based on frost points reported at various elevations. Mixing ratios given on this chart correspond to the indicated frost points only when the pressure at a given elevation is the same as in the U. S. Extension to the ICAO Standard Atmosphere, and hence they only approximate actual mixing ratios. Identifications of the soundings are: A, Thule, Greenland, 8 Apr 60; B, Washington, D. C., 8 Apr 60; C, Churchill, Canada, 4 Nov 58; M, Camp Ripley, Minn., 1 Jul 49; P, Camp Ripley, Minn., 26 Aug 49; V, Washington, D. C., 8 Feb 60; W, Washington, D. C., 27 Jun 60.

TABLE 1. Frost Points and Mixing Ratios for Typical American and British Hygrometric Ascents at Level of 14 to 15 Kilometers

Location	Lat	Long	Date	Frost Point, °C	Mixing ratio, mg
Thule, Greenland	77N	94W	8 Apr 60	-67	17.8*
Ft. Churchill, Canada	59N	94W	4 Nov 58	-60	47.4*
Ft. Churchill, Canada	59N	94W	8 Apr 60	-85	1.0*
Camp Ripley, Minn.	47N	93W	1 Jul 49	-77	3.6
Camp Ripley, Minn.	47N	93W	26 Aug 49	-59	55.7
Washington, D. C.	39N	77W	8 Apr 60	-81	1.9
Washington, D. C.	39N	77W	27 Jun 60	-77	4.3
Washington, D. C.	39N	77W	8 Feb 60	-70	11.6
Denver, Colo.	40N	105W	28 Apr 59	-60	46.3
St. Louis, Mo.	39N	90W	7 Jan 50	-59	55.2
Norwegian Sea	67N	3W	24 Jul 56	-82	2.1
S. England	51N	1W	22 Dec 54	-83	1.7
S. England	51N	1W	6 Jul 54	-82	2.0
S. England	51N	0	22 Nov 55	-83	1.7
S. England	51N	1W	24 Feb 54	-82	1.6
S. England	51N	1W	5 Jul 54	-82	1.6
W. Mediterranean	40N	5E	24 Jul 56	-82	1.8
W. Mediterranean	41N	5E	4 Mar 57	-80	2.4
W. Mediterranean	42N	4E	19 Jul 56	-82	1.8
Idris, Tripoli	33N	13E	30 May 56	-83	1.5

* For 142 mb.

between ascent and descent but a spatial differential as well. Ascent and descent can be in exact agreement when the large-scale weather situation is static; selection of only such cases provides a biased sample.

So confused is the subject of stratospheric humidity that at least three investigators [Gutnick, 1960; Hales, Williams, and Henderson, 1960; Greenfield and Kellogg, 1960] abandoned efforts to estimate it for their studies and relied on upper and lower bounds (moist stratosphere and dry sky), hoping that the actual moisture distribution would be bracketed.

Study of the current status of theory and observation concerning stratospheric humidity raises several questions, none of which can yet be answered: (a) Why do British data consistently indicate a dry sky, while American ascents indicate that the moisture content of the lower stratosphere is extremely variable? (b) What mechanism can explain the increase of mixing ratios with height, shown by most American ascents? (c) Can a dynamically sound model be constructed to fit the available observations?

Under the Dobson-Brewer meridional model, a dry lower stratosphere should be observed at all but high latitudes; this is not substantiated by the American data. The model also does not

explain how the tropospheric air penetrates well-developed, strong, and persistent tropopause with any regularity. Although many such mechanisms have been suggested, none yet been proved [Machta, 1959].

Observation of relatively moist air at 27° over England by Barclay and others also conflicts with the Dobson-Brewer model, but could be reconciled by further postulates. For example, moist tropospheric air may be injected into the stratosphere by disturbances which penetrate the weak polar tropopause. This relatively moist air spreads over the dry layer in lower latitudes by turbulent mixing and diffusion. Such a model could fit the British data, but not the American observations, which show variable moisture in the lower stratosphere.

Conclusions. Lack of a tenable explanation of the various conflicting reports on stratospheric humidity casts doubt on the accuracy of past all of the data. If the validity of the data can be questioned, the first and most obvious statistical points are the instrumentation and/or techniques. Although virtually every scientist who has studied stratospheric moisture has called for additional data, making more human ascents will only amplify the confusion. If the instrumentation and/or techniques are in question (as might well be the case), then formulate

hypothesis based on such data will lead to erroneous conclusions.

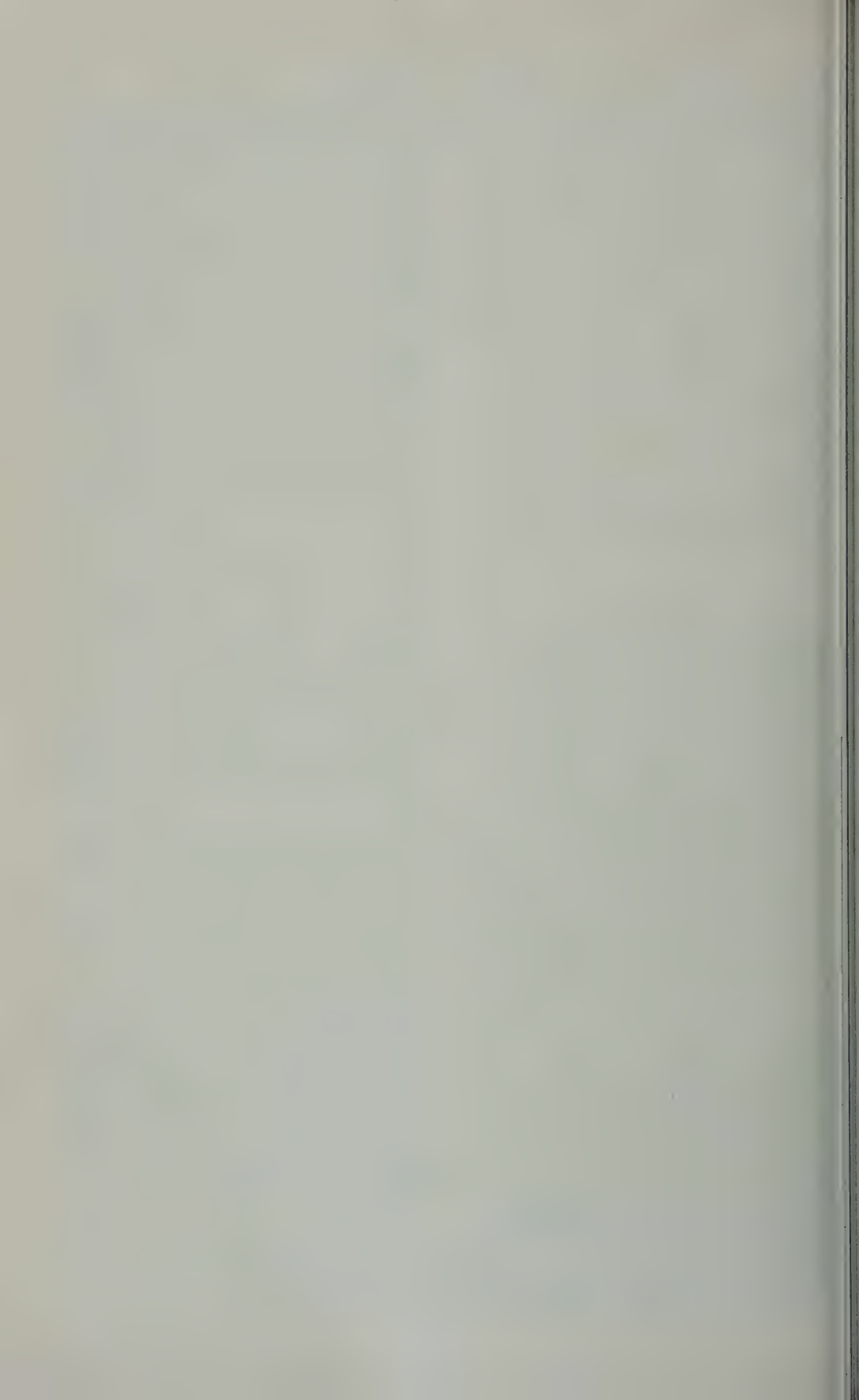
Re-evaluation of the instrumentation is the most logical step to take. This will involve rigorous laboratory and field tests of the instruments and techniques used to gather the stratospheric humidity data. Such tests may reveal that one or more of the devices is in error. The error gathered by this instrument or method may be eliminated or, if possible, corrected, giving enough accurate data for a tenable hypothesis to be formulated. Such an evaluation will be expensive and beset with many technical difficulties. But it must be done, if we are to eliminate instrumental error as a possible cause of the present confusion about stratospheric humidity.

Acknowledgments. Encouragement and guidance in the preparation of this report were given by Herman Sissenwine, Irving Gringorten, and especially Arnold Court, who edited and titled it.

REFERENCES

- Bay, F. R., M. J. Elliott, P. Goldsmith, and J. V. Kerley, A direct measurement of the humidity in the stratosphere using a cooled-vapor trap, *Quart. J. Roy. Meteorol. Soc.*, **86**, 259-264, 1960.
- Brett, E. W., L. R. Herndon, and H. J. Carter, Humidity measurements of the distribution of water vapor in the stratosphere, *Tellus*, **2**(4), 302-311, 1950.
- Chapman, A. W., Evidence for a world circulation provided by the measurements of helium and water vapor distribution in the stratosphere, *Quart. J. Roy. Meteorol. Soc.*, **75**, 351-363, 1949.
- Chapman, A. W., Ozone concentration measurements made in an aircraft in N. Norway, *MRP 946*, Meteorological Research Committee, London, 1955.
- Chapman, G. M. B., Origin and distribution of the heavy atomic molecules in the atmosphere, *Proc. Roy. Soc., London*, **236**(1205), 187-192, 1956.
- Cole, D. M., D. G. Murcray, C. C. Shaw, and R. J. Arnold, Near infrared solar measurements by balloon to altitudes of 100,000 feet, *J. Opt. Soc. Am.*, **48**, 1010-1016, 1958.
- Goldsmith, P., Some aircraft and surface meteorological observations made at Khartoum, *Meteorol. Mag.*, **23**, 329-334, 1954.
- Hartfield, S. M., and W. W. Kellogg, Calculations of atmospheric infrared radiation as seen from a meteorological satellite, *J. Meteorol.*, **19**(3), 287-290, 1960.
- Hick, M., An estimate of precipitable water along high altitude ray paths, *Air Force Surveys in Geophysics 120*, Air Force Cambridge Research Laboratories, Bedford, Mass., 1960.
- Helliwell, J. V., T. L. Williams, and D. Henderson, Calculation of infrared radiative flux emission plus atmosphere at various levels high above the earth, *Final Report for Geophysics Research Directorate, contract AF 19(604)-2418, AFCRL*, 1960.
- Helliwell, N. C., Airborne measurements of the latitudinal variation of frost point, temperature and wind, *Sci. Paper 1*, Air Ministry Meteorological Office, London, 1960.
- Helliwell, N. C., and J. K. Mackenzie, Observations of humidity, temperature and wind at Idris 23rd May-2nd June, 1956, *MRP 1024*, Meteorological Research Committee, London, 1957.
- Helliwell, N. C., J. K. Mackenzie, and M. J. Kerley, Further observations of humidity up to 50,000 feet, made from an aircraft of the Meteorological Research Flight in 1955, *MRP 976*, Meteorological Research Committee, London, 1956.
- Helliwell, N. C., J. K. Mackenzie, and M. J. Kerley, Some further observations from aircraft of frost point and temperature up to 50,000 feet, *Quart. J. Roy. Meteorol. Soc.*, **83**, 257-262, 1957.
- Houghton, J., and J. S. Seeley, Spectrographic observations of the water vapor content of the stratosphere, *Quart. J. Roy. Meteorol. Soc.*, **86**, 357-370, 1960.
- Kerley, N. J., High altitude observations between the United Kingdom and Nairobi, *Meteorol. Mag.*, **90**(1062), 3-17, 1961.
- Machta, L., Transport in the stratosphere and through the tropopause, *Advances in Geophysics*, **6**, Academic Press, New York and London, 273-288, 1959.
- Masterbrook, H. J., and J. E. Dinger, Distribution of water vapor in the stratosphere, *J. Geophys. Research*, **66**, 1437-1444, 1961.
- Murcray, D. G., F. H. Murcray, W. J. Williams, and F. E. Leslie, Water vapor distribution above 90,000 feet, *Sci. Rep. 5*, for *Geophysics Research Directorate, contract AF 19(604)-2069, AFCRL*, 1960.
- Murgatroyd, R. J., P. Goldsmith, and W. E. H. Hollings, An interim report on measurements of humidity from aircraft to heights of about 50,000 feet over southern England, *MRP 877*, Meteorological Research Committee, London, 1954.
- Murgatroyd, R. J., P. Goldsmith, and W. E. H. Hollings, Some recent measurements of humidity from aircraft up to heights of about 50,000 feet over southern England, *Quart. J. Roy. Meteorol. Soc.*, **81**, 533-537, 1955.
- Sissenwine, N., and M. Gutnick, Precipitable water along high altitude ray paths, *Proc. Infrared Information Symposium*, **5**(3), 5-11, 1960.
- Tucker, G. B., An analysis of humidity measurements in the upper troposphere and lower stratosphere over southern England, *MRP 1052*, Meteorological Research Committee, London, 1957.
- Yarnell, J., and R. M. Goody, Infrared solar spectroscopy in a high-altitude aircraft, *J. Sci. Instr.*, **29**, 352-357, 1952.

(Manuscript received April 10, 1961; revised June 14, 1961.)



Lava-Sea-Air Contact Areas as Sources of Sea-Salt Particles in the Atmosphere¹

A. H. WOODCOCK AND A. T. SPENCER

Woods Hole Oceanographic Institution, Woods Hole, Massachusetts

Abstract. One theory of individual raindrop formation requires a giant hygroscopic particle ($\geq 10^{-12}$ gram) as an initiating nucleus. For this and other reasons, natural sources of these nuclei are of interest to meteorologists. The sea surface is normally a source of relatively low numbers of these nuclei. Enormous numbers are present, however, in the steam clouds arising when molten lava encounters sea water. Observations of the number and the size of the steam-cloud particles or nuclei are given. Samples of these particles were taken from aircraft within the cloud during the 1960 Kilauea eruption in Hawaii. Other samples were taken from experimental clouds formed in the laboratory. Microscope and flame-photometric methods of analysis showed the particles to be, in major part, sea salt. Size distribution studies reveal that the flash boiling, resulting from lava-sea-air contact, produces salt nuclei of a size range in large part like that normally arising from the sea. The rate of production of these nuclei, expressed as grams per unit of surface, is from 1 to 6 million times greater from steaming lava than from the average sea surface.

I. INTRODUCTION

Recent eruptions of the Kilauea volcano on the island of Hawaii have produced large steam clouds in areas where the lava, at a temperature about 1000°C [Eaton and Murata, 1960], flowed into the sea (Fig. 1). This flow continued for several weeks, and marked changes in visibility occurred 320 km away at Honolulu on the island of Oahu. Also, aircraft and surface observers observed dense haze at much greater distances over the North Pacific.

Simpson [1950] reported the formation of the haze associated with the 1950 eruption. During the 1960 eruption, S. Price (personal communication, 1960) noted that haze formation followed the entry of the lava streams into the sea, and he supposed that the haze-producing particles were largely sea salts which were thrown into the air by the violent vaporization or flash boiling of the sea in contact with the lava. The present study confirms Price's idea that the steam-cloud particles are largely sea salts.

Recent work has emphasized the meteorological and geochemical importance of large airborne salt nuclei produced by the sea [Blanchard, 1961; Eksson, 1959; Woodcock and Blanchard, 1955; Woodcock, 1952, 1960]. The role of this marine aerosol in the formation of the large cloud droplets essential to accretional growth of raindrops is becoming recognized [Mason, 1957]. It is

pertinent to expand our knowledge of the sea as a source of atmospheric salts to include those unusual areas and times at which the heat of volcanic activity forces salts into the air at an extraordinary rate.

The role of explosive volcanoes in occasionally adding vast quantities of finely divided material to the atmosphere is well known [Symons, 1888]. Eruptions of the nonexplosive shield-type volcanoes of Hawaii can also add great quantities of finely divided (hygroscopic) material to the air by the relatively slow but long-sustained flow of lava into the sea. It is supposed that other volcanic activity at or near sea level, such as that of Myojin-Sho near Japan [Dietz and Sheehy, 1954] will be especially productive of atomized sea salt.

The purpose of this paper is to present data on the size distributions of the salt particles added to the air in the steam clouds and to make estimates of the rates of production. These data were obtained over Hawaiian lava flows and in experimentally produced clouds in the laboratory. In the concluding section, estimates are made of the total quantities of particles added to the atmosphere by a major lava flow, and some discussion is given of physical factors which may alter these quantities.

2. NATURAL STEAM-CLOUD PARTICLES

In this section a brief discussion is presented of the techniques used to sample the steam-cloud

¹Contribution 1171 of the Woods Hole Oceanographic Institution.



A



B

Fig. 1. Steam clouds formed as lava from the 1960 eruption of Kilauea enters the sea in the Puna district, island of Hawaii. Scale in picture *B*, may be judged from height of lava near sea, which is about four meters. Photograph *A* by Jack Pales, U.S.W.B., and photograph *B* by Robert Haugen of the National Park Service, Hawaii.

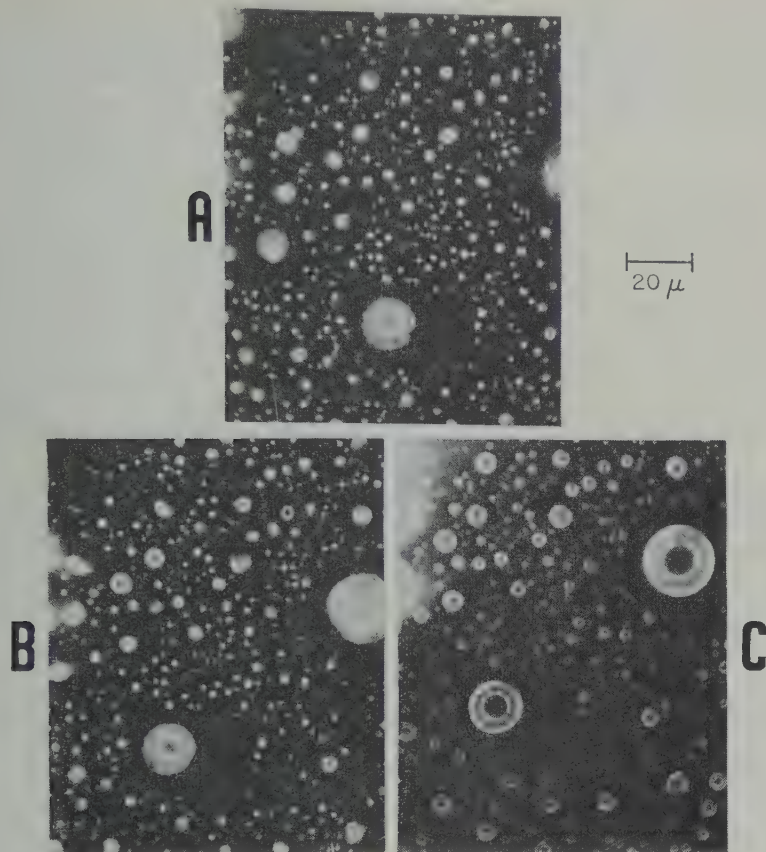


Fig. 2. Photomicrographs of steam-cloud particles on surface of sampling slide which was exposed to different relative humidities. See text for discussion.

particles and of the physical-chemical methods which show that they are largely sea salt.

Fifteen samples of the steam-cloud particles were collected on two sampling days by causing particles to impinge upon prepared 1×20 cm glass slides. These slides were exposed from aircraft by means of a device previously described by Woodcock [1952, p. 208]. Flight speeds of 40 m sec^{-1} were used, and the samples were taken in the smaller steam clouds, such as those shown on the left in Figure 1B. Droplets and moist particles of certain size ranges impinge upon and adhere to these slides. A hydrophobic Dri-film coating causes droplets on them to form hemispheres [Woodcock and Gifford, 1949], thus simplifying volume estimates from diameter measurements.

Figure 2 shows photomicrographs of these steam-cloud particles. They contain cubic crys-

tals at relative humidities of 71 per cent and less,² as shown in pictures B and A. They change phase at 74 per cent and are completely liquified at 77 per cent. These are properties of sea-salt particles [Twomey, 1953, 1954].

Examination with a microscope of many hundreds of droplets, such as those in Figure 2C, revealed that there were practically no insoluble particles present on the slides. Since soil, sand, and other small insoluble particles are readily deposited on the slides, the absence of these materials suggests that small particles of lava

² Immersion of the slides in water-vapor-saturated air over a 35 per cent solution of sulfuric acid and water (RH 71%) causes the amorphous mass of sodium chloride crystals usually formed to slowly recrystallize into larger cubic crystals of a lower-energy state.

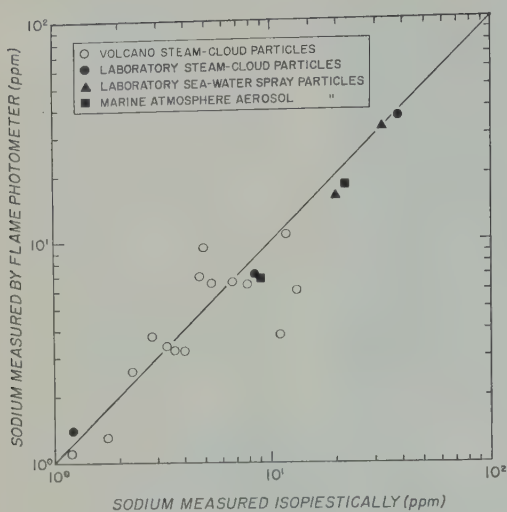


Fig. 3. Comparison of isopiestic and flame-spectrophotometer measurements of sodium in steam-cloud and other particles. Most of the scatter of the observations about the 45° line is attributed to the normal error of the two methods. See text.

or other materials are not carried upward very far in these clouds.

From diameter measurements of the hemispheric droplets, at a relative humidity of 91 per cent, and from the vapor pressure of concentrated sea water, one can compute the weight of sea salt W (grams) in each droplet.

$$W = V d C \quad (1)$$

where

V = volume of hemispheric droplets (μ^3).

d = density of droplets [*Higashi, Nakamura, and Hara, 1931*] (g ml^{-1}).

C = concentration of salts in droplets by weight fraction [*Arons and Kientzler, 1954*].

The weights of salt in a great many droplets on each of the slides were thus determined. Estimates were then made of the total sea salt in all the droplets on the slides, the number of which usually greatly exceed those actually measured under the microscope. Subsequently all the salts on each slide were dissolved in from 50 to 200 mg of distilled water, and a sodium analysis was made with a flame spectrophotometer. Thus it was possible to compare the concentrations of sodium in these solutions, which were computed from the isopiastically

determined weights of sea salt, to the sodium concentrations photometrically measured. The small-volume samples were readily measured with a modified Beckman 4040 atomizer burner.

Figure 3 shows a comparison of the sodium amounts derived by the two methods, isopiastically determined sodium concentration being derived from the product of the sea-salt weights and the ratio of the sodium in them to the total salts. Three of the 22 observations fall outside the range of error expected from combined errors of the two methods. The unusual observations may be due to a lowering of the surface tension by contaminating film resulting in a flattening of the larger droplets and an overestimate of their size and sodium content by the isopiestic technique. They may also result from the occasional failure of larger particles, on that fraction of the slide used for measurement under the microscope, to be representative of those on the slide as a whole.

Among 20 test measurements of the error of the flame photometer in the range of concentration concerned, 90 per cent showed an error $< \pm 10$ per cent. All fell within ± 15 per cent of the test solutions. The error of the isopiestic method, as used here, is about ± 20 per cent.

From the agreement revealed in Figure 3, and from the other properties of the crystal particles previously mentioned, it is concluded that the steam-cloud particles are in major part sea salt.

3. STEAM-CLOUD SALT PARTICLES COMPARED WITH SALT NUCLEI IN MARINE AIR

As an aid in thinking about the meteorological role of volcano steam-cloud salt particles, it is useful to derive their spatial distributions and compare them with those of salt particles normally found in the air over the sea.

Figure 4, curve A, shows an average steam-cloud particle distribution curve, compared with the normal range of average concentrations of maritime air under wind forces 3, 5, 7, and 9. Curve A is based upon micrometer measurements of the numbers and sizes of many hundreds of individual particles. These measurements are then readily converted to spatial distributions in the air, using the observed aircraft speeds, slide exposure durations and areas.

The potency of the lava-sea-air contact

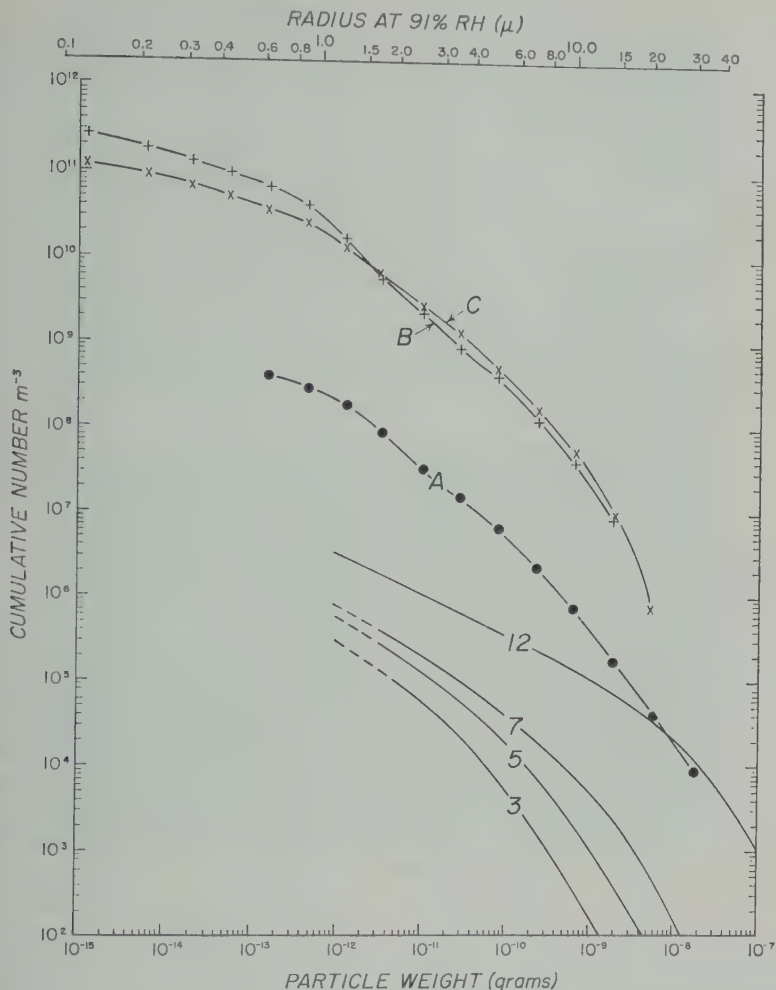


Fig. 4. Cumulative steam-cloud particle distribution curves (A, B, and C) compared with averagesalt-particle distributions found at the 300- to 600-meter altitudes over the sea during winds of forces 3, 5, 7 and 12. Symbols representing particle counts are plotted at the lower limiting particle weight of each weight range category. Curve A represents average particle concentrations in natural rain clouds, and B and C give concentrations in artificially formed clouds in the laboratory. Upper scale gives particle size as measured under the microscope and converted to equivalent spherical radius.

source of salt particles is obvious. It is clear that this zone produces especially great numbers of smaller particles and that the range is comparable to that of particles normally derived from the sea throughout most of the size range.

It is probable, however, that the concentration of particles in the small steam clouds ($\approx 10^8$ particles $\text{m}^{-3} \geq 3 \times 10^{-13}$ g) was greater than that which might have been found in large clouds present during a main advance

of lava into the sea, such as that shown at the top left side of Figure 1. Note the comparatively dense character of this cloud. This impression is also supported by the reported observation, during the great 1950 flow of lava into the sea, of about 10^{-9} gram of salt per cubic centimeter of air in Honolulu [Simpson, 1950], which was 320 km downwind from the source area. It seems that the salt particle concentration near the steam-cloud source in 1950 may have been much greater than the quantity reported here (i.e.

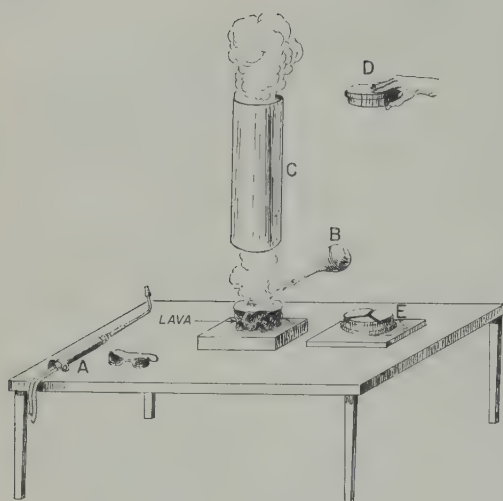


Fig. 5. Sketch illustrating method used to trap steam-cloud particles produced in the laboratory. Cylinder *C* is 44 cm long and 14 cm in diameter.

4×10^{-9} g cm $^{-3}$), which was measured only a fraction of a kilometer from the lava-sea-air contact zone.

It will be demonstrated later than there were probably many more small particles present in the Hawaiian steam cloud than those represented in curve *A*, Figure 4. These small particles add little, however, to the total mass per unit volume of air, which is discussed in the preceding paragraph. Langmuir has shown that the collection efficiency *E*, defined as the ratio of the actual deposition per unit surface on the slide or 'ribbon' to the total possible deposition under the sampling conditions, is negligible for the 1-mm-wide glass slides, at an air speed of 40 m sec $^{-1}$ and for particles of density 1 and smaller than about $0.5\text{-}\mu$ radius [Langmuir and Blodgett, 1945; Woodcock and Gifford, 1949]. Correction for *E* cannot be applied to these steam-cloud particle samples because particle size at sampling time is not known. It can be stated, however, from Langmuir's theoretical study of impact deposition on 'ribbons' [Langmuir and Blodgett, 1945], that *E* is at least 0.8 for particles $> 10^{-11}$ g and at least 0.5 for those $> 10^{-12}$ g. Failure to apply this *E* correction to the steam-cloud samples is not a serious handicap, however, in the present study.

A further, more serious, question about the natural steam-cloud samples concerns the degree to which droplet or particle coalescence, at the

time of impact on the slides, may have altered the particle size distribution. A reasonable 'covering fraction' on the slides [see Woodcock and Gifford, 1949, p. 189, and Langmuir, pp. 59-70] could not be established since there was no way to know the particle sizes in the clouds and to adjust the sampling distance accordingly. Thus the size distribution should be regarded as a best approximation. The comparable nature of the size distribution on curves *A*, *B*, and *C* of Figure 4 causes the authors to think that coalescence has not altered the distribution of the natural cloud particles.

4. ESTIMATE OF RATE AT WHICH THE AEROSOL WAS PRODUCED WHERE LAVA ENTERED SEA

Cine photographs of a small steam cloud similar to that in which samples were collected showed that the diameter of the plume at the sampling level was much greater than the diameter close to the source. A rough value for the dilution factor of particles was derived from measurements on photographs of the diameters of the smoke plumes near the sea surface at altitude and close to the source area. From airplane speed and the time required to travel through the steam clouds normal to the axes, an estimate of the diameter at sea surface altitude was made. With this diameter and relative sizes derived from the steam cloud photographs, the plume size near the source was derived.

Using this estimate of the enlargement of the steam-cloud plume from source to sampling level and a rate of rise derived from the angle of the plume in a known wind, a particle emission rate *P* of about 1.4×10^6 cm 3 sec $^{-1}$ was determined for those $\geq 3 \times 10^{-12}$ g.

$$P = CRV$$

where

C = number concentration of particles $\geq 3 \times 10^{-12}$ g at sampling level (cm $^{-3}$).

V = rate of rise of cloud normal to cross section (about 600 cm sec $^{-1}$).

R = ratio of cloud cross section at sampling level (3220 m 2) to that near source (129 m 2).

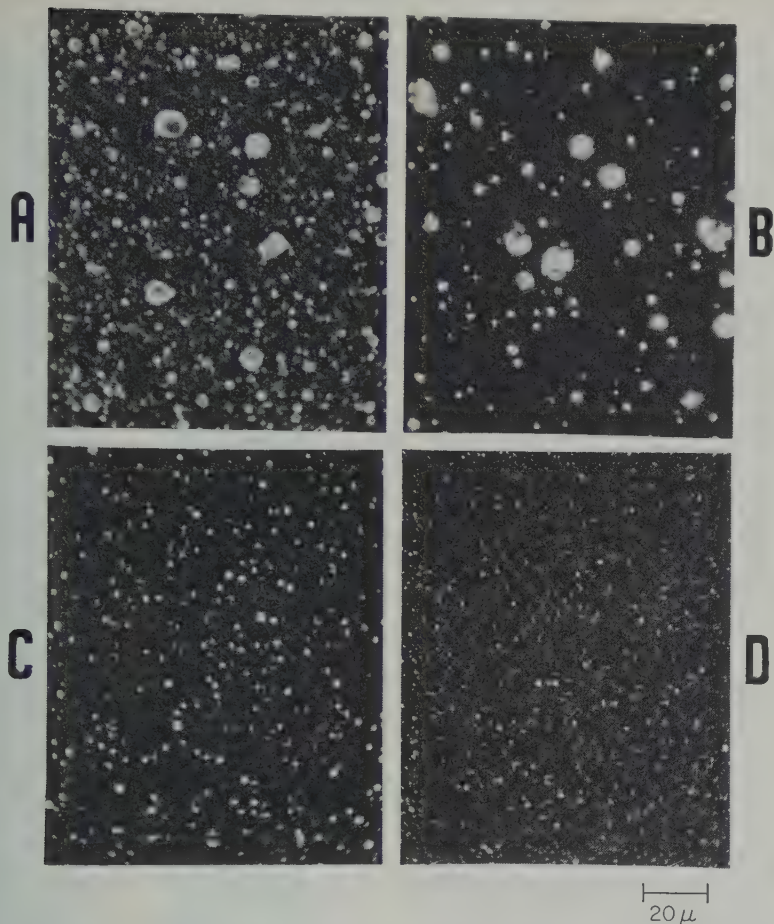


Fig. 6 Photomicrographs of salt crystals on glass slides exposed to fallout from laboratory steam cloud in sedimentation chamber. A shows excess 'covering fraction' or over-sampling from exposure to entire fallout. On slides B, C, and D, samples were taken from time 0 to 200, 200 to 1170, and 1170 to 7×10^4 seconds, respectively, thus avoiding overlapping on any one slide. See text.

of the weights of these particles gave a reduction rate of $4.5 \times 10^{-5} \text{ g cm}^{-2} \text{ sec}^{-1}$. To test the above estimates and to learn if smaller particles may also have been present, it was decided to make miniature steam clouds in the laboratory. These tests are described in the following section.

PARTICLES FROM ARTIFICIAL STEAM CLOUDS

A sample of lava from the Kilauea eruption of 1912 which D. H. Richter of the Hawaiian Volcano Observatory kindly sent to us was placed in a molten red-yellow state with an oxyacetylene torch, as suggested in Figure 5. The incandescence had penetrated to a

depth of about 1 cm, the flame was removed and the molten surface splashed with sea water B.³ The steam cloud which arose from the lava was caused to flow through the cylindrical sedimentation chamber C. When steam was visibly entering the base of the open-ended cylinder and passing out at the top, a cover D was placed over the top, the lava removed, and the cylinder lowered over a prepared base E

³ The lava in this case presented a relatively smooth surface to the water, a type of surface characteristic of pahoehoe lavas. It is thought that the porous surface of aa lavas would produce a different result when coated with sea water. (see end of section 8.)

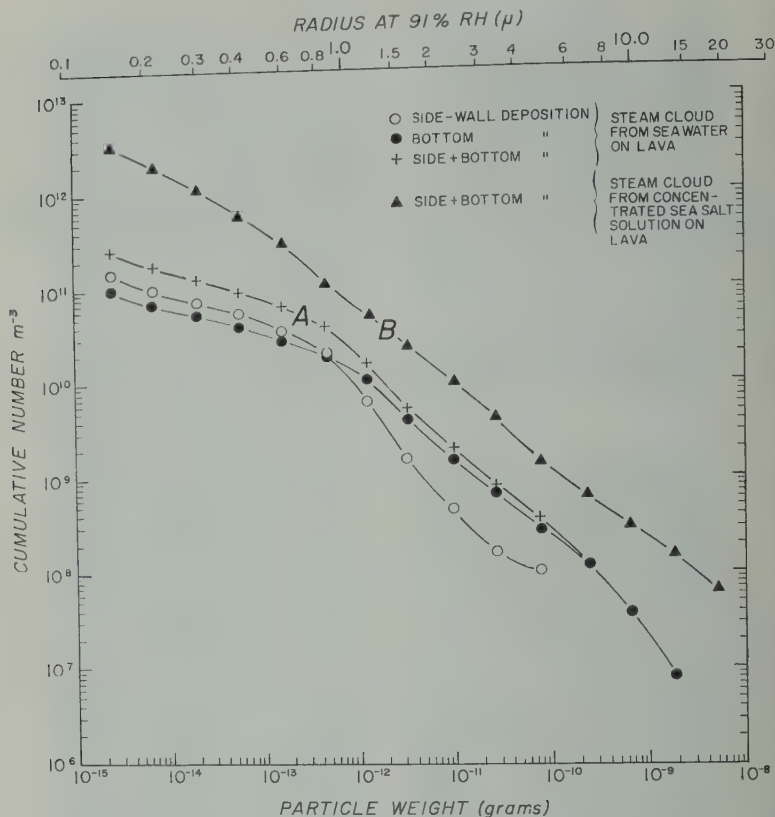


Fig. 7. Size distribution and relative deposition curves, representing steam-cloud particles deposited on the side walls and bottom of the sedimentation or fallout chamber. Symbols representing particle counts are plotted at the lower limiting particle weight of each weight category.

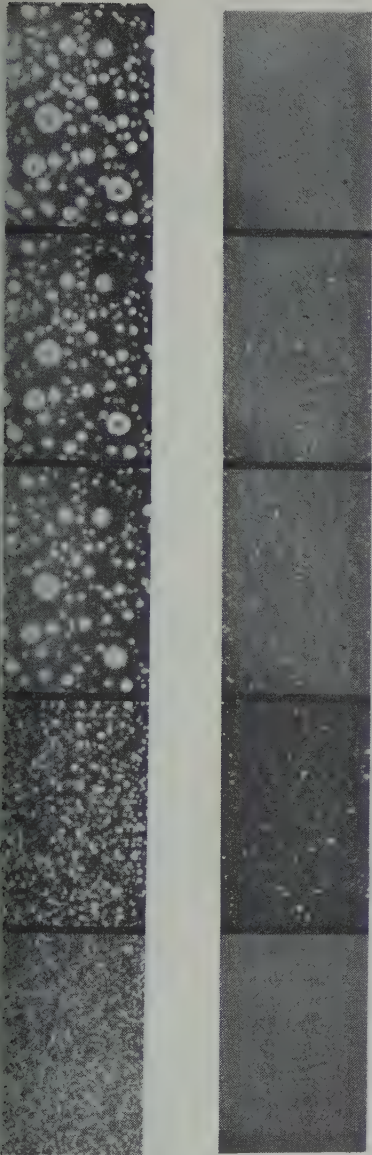
containing a glass-slide sampling surface. All joints were effectively sealed with masking tape and the chamber set aside to await fallout of the particles.

The chamber was found to contain so many particles that the sampling slide was excessively covered, as illustrated in Figure 6A. This excess can result in particle coalescence on the slides as the crystals become relatively large droplets at the high relative humidities required to measure them isopiesticly. Later samples were taken on several slides (see Fig. 6B, C, and D), thus avoiding this problem of an excess covering.

In Figure 4, curves B and C, it can be seen that the number of particles in the air close to the steaming lava surface in the laboratory is about two orders of magnitude greater throughout the range of sizes than the average of those sampled at about 100 m over the Hawaiian lava flow (curve A, Fig. 4). It will be shown below,

however, that the actual particle production rate at the lava surface in the laboratory is that estimated above from the natural clouds.

Metal foil was used as a lining for the surface of the sedimentation chamber to eliminate possible local charge effects which might cause side-wall deposition [See Cadle, 1957]. Later examination with microscope of sections of this foil cut from the chamber revealed, however, that deposition of particles seemed to occur. The steam-cloud experiment was repeated with several slides attached to the walls of the chamber as well as to the bottom. As suspected, it was found that many of the smaller particles were deposited on the walls. Figure 7 shows the relative magnitude of this side-wall deposition, and Figure 8 shows photomicrographs of the particles deposited on the side-wall slide surfaces. It can be seen in



H 50μ

Photomicrographs of laboratory steam-cloud particles, from saturated sea-salt solution on slides deposited on glass sampling slides in sedimentation chambers. See text.

all deposition is greater for particles at 5×10^{-13} g. It is supposed that this position is due to thermal effects. Figure 8 further illustrates the sedimentation for sampling salt crystals. The photo-

graphs show the steam-cloud particles, from a saturated sea-salt solution on hot lava, deposited on slides taken from the chamber. In frames 11, 12, and 13, the microscope was focused on the tops of the largest, medium-sized, and smallest cubic crystals, respectively, which were deposited by fallout on the bottom during the first 120 seconds. The purpose was to show something of the range of size of the cubic crystals of sodium chloride, which were clearly resolved under the microscope. Frames 14, 15 and 16 show deposition on other slides during successive periods of 900 seconds (i.e. 120 to 1020 seconds from time 0), 75,000 seconds, and 88,000 seconds, respectively. Side-wall deposition is shown on slides represented by frames 17, 18, and 19, which were attached near the top, middle, and bottom of the inner surface. Frame 20 shows the surface of the 'control' slide, which was exposed for 75,000 seconds to fallout from room air sealed in a separate sedimentation chamber.

A small Scholz condensation nucleus counter [Landsberg, 1938, p. 192] was used to demonstrate that the direct counts with the microscope of the individual particles falling in the sedimentation chamber actually represented most of the condensation nuclei present. Table 1 shows the results. For present purposes, the differences between the counts obtained by the two methods are not significant.

Many of the smaller particles, which are mere points of light on Figure 6D, were examined with an electron microscope. All particles were found to contain crystalline material, and cubic sodium-chloride crystals were visible in most of them. The smallest crystal seen was about 0.1μ in diameter.

TABLE 1. Counts of Condensation Nuclei Made in Room Air and in the Laboratory Steam Cloud

Test No.	Scholz Counter*		Direct Count Steam-Cloud Particles with Microscope, † $N \text{ cm}^{-3}$
	Room Air, $N \text{ cm}^{-3}$	Steam Cloud, $N \text{ cm}^{-3}$	
2	0.02×10^5	3.2×10^5	2.7×10^5
3	0.27×10^5	11×10^5	33×10^5

* An average, in each case, of many individual counts.

† Number computed from counts of more than 1800 crystals.

TABLE 2. Comparison of the Rates of Production of Salt Particles at the Surface of Molten Lava on the Puna Coast and in the Laboratory

	Puna Coast Steam-Cloud Particles, $\geq 3 \times 10^{-12}$ g (1)	Laboratory Steam Clouds		
		Sea Water Spray		Saturated Sea Salt
		Particles $\geq 3 \times 10^{-12}$ g (2)	All Particles (3)	All Particles (4)
Salt production, grams $\text{cm}^{-2} \text{sec}^{-1}$	4.5×10^{-5}	7×10^{-5}	10×10^{-5}	45×10^{-5}
Particle production, number $\text{cm}^{-2} \text{sec}^{-1}$	1.4×10^6	2×10^6	100×10^6	745×10^6

Some difficulty was encountered with the nucleus counts due apparently to particle deposition on the walls of the short rubber tube leading from the cloud-filled chamber to the intake valve of the Scholz counter. This was solved by inserting a syringe needle through the tube wall at the counter intake orifice and, immediately before each sampling, drawing through the tube a volume of cloud air equal to several times the total inside volume of the tube. This step assured a minimum time for deposition on the inner walls of the tube.

On each occasion when steam clouds were sampled, background counts of condensation nuclei in the laboratory room air were also made. These are given in Table 1, with the condensation nucleus and salt crystal counts.

From the rate of rise, horizontal dimensions, and salt concentration of the miniature steam clouds, and from the surface area of the molten lava, estimates were made of the rate of production of salt particles P .

$$P = CAV/a \tag{3}$$

where

- C = particle number cm^{-3} .
- A = cross-section area of sedimentation chamber (154 cm^2).
- V = rate of ascent of cloud (22 cm sec^{-1}).
- a = area of hot lava surface (15 cm^2).

The average rate of production of particles was about $2 \times 10^6 \text{ cm}^{-2} \text{sec}^{-1}$. This same expression can be similarly used to derive a total weight of sea salt of $7 \times 10^{-5} \text{ g cm}^{-2} \text{sec}^{-1}$, when limited to particles having a weight $\geq 3 \times 10^{-12} \text{ g}$.

Table 2 shows a comparison of particle production by the laboratory steam cloud and

by the Hawaiian clouds which were sampled. The quantities in columns 1 and 2 are regarded as an indication that the flash boiling process observed in the formation of the laboratory clouds also occurred in the formation of natural clouds from the Hawaiian lava fields. From this result it is also concluded as probable that the numerous smaller particles found in the laboratory clouds were also present in the natural steam clouds, though unsampled due to the limitations of the method used over the lava fields.

From the rate at which salt was driven from the lava surface by this flash boiling of the sea water, a rough estimate of the evaporation E and heat utilization H were also made. They are expressed as follows:

$$E = W/S$$

where

- E = sea water evaporated ($\text{g cm}^{-2} \text{sec}^{-1}$)
- W = rate at which sea salt forced into air ($\text{g cm}^{-2} \text{sec}^{-1}$).
- S = salinity of sea water, expressed as a fraction.

$$H = Eh$$

where

- H = calories used $\text{cm}^{-2} \text{sec}^{-1}$.
- E = sea water evaporated ($\text{g cm}^{-2} \text{sec}^{-1}$)
- h = heat of vaporization of sea water (g^{-1}).

By utilizing equations (3), (4), and (5) a quantity of about $2 \text{ cal cm}^{-2} \text{sec}^{-1}$ was computed for the rate of consumption of heat during a few seconds of flash boiling of the water.

lava. The physics handbook indicates a thermal conductivity of basalt of about 0.005 cal/cm² sec/°C. Thus a brief heat flow of 2 cal/sec/cm² can be expected with a temperature gradient of about 1000°C in somewhat less than 1-cm thickness of lava. Some additional heat is also obtainable from the latent heat of crystallization of lava.

In making the above estimates of the rates at which water is vaporized, it is assumed that all salts are driven off into the air with the vapor. It is observed, however, that some salts remain on the surface of our laboratory lava after it is dried and relatively cooled. The extent to which a salt residue such as this remains under natural conditions is unknown. The above heat-flow estimate is therefore somewhat less than the actual amount.

Most of the water disappears and a salt residue is left, the particle production rate is greatly increased, if the lava remains molten. This was tested in the laboratory by spraying a saturated solution of sea salts on molten lava and trapping a portion of the white cloud particles which arose. Figure 8, curve B, shows the comparative numbers and weights of salt particles found; Table 2, column 4, shows the particle production rate; and Figure 8 shows photomicrographs of the particles formed. It is thought that this increase in the number of particles may result from the availability of heat for the fractionation process, which is unknown.

As for the apparent vaporization of all water, the vapors rising from the salt-coated hot lava, when inhaled, are occasionally irritating to the throat. D. E. Carritt, chemist of this laboratory, states that this irritant is probably sulfuric acid (and/or) H₂SO₄, resulting from the breakdown of the chlorides and sulfates.

Eaton has reported (personal communication) that rains falling from or through steam clouds were occasionally acidic, causing a burning sensation to the skin and eyes. These observations of the irritating character of the vapors suggests that in some cases a part of the lava remains hot enough, after contact with the sea, to break down salts remaining on the water is driven off.

An interesting fact that sea water droplets, when splashed or sprayed on molten lava, tend to remain on the surface during vaporization and

do not form the almost spherical, dancing droplets so characteristic of water on very hot metals. This difference in performance is thought to be due to the low thermal conductivity of lava, which is only about one-twentieth of that of steel and about a hundredth of that of copper.

It is of interest to meteorologists and others to compare the rates of production of salt particles by the hot lava in contact with the sea to that of the sea surface as a whole. Blanchard [1961] recently computed the average annual world fallout of sea salt to be 9×10^9 metric tons. This estimate is based upon measurements of salt found as particles at cloud levels in the atmosphere and salt in solution in falling rain. From Blanchard's estimate it is found that 1 m² of sea surface produces about 0.8 μg of salt per second, which reaches cloud-level altitudes. The total surface area of the oceans and seas assumed here is 3.6×10^8 km².

From the field and laboratory results given here it is computed that a square meter of lava produces from 0.7 to 4.5 grams of finely divided salt particles per second, or an amount equal to that produced on the average by about one to six million square meters of sea surface. In other words, the hot lava produces airborne salt from one to six million times faster per unit area than the average of the seas. Expressed in terms of particle number ($\geq 3 \times 10^{-12}$ gm), the range of rates of production, extended over the duration of the 1960 flow, is such that 2 to 7 km² of hot lava surface could double the salt-aerosol population within the first kilometer of the earth's atmosphere. These estimates emphasize again the relatively great particle productivity of hot lava at sea level.

6. MAJOR FACTORS AFFECTING DENSITY OF STEAM CLOUDS

Our experience with natural and artificial steam clouds has caused the authors to think that large differences should be found in the concentration of particles in these clouds. The physical factors producing these differences may be grouped under two headings: the rate and mechanics of availability of heat for the flash boiling of sea water and the manner in which the sea water is brought into contact with the lava.

Factors altering heat availability. Extreme examples of differences in the mechanics whereby

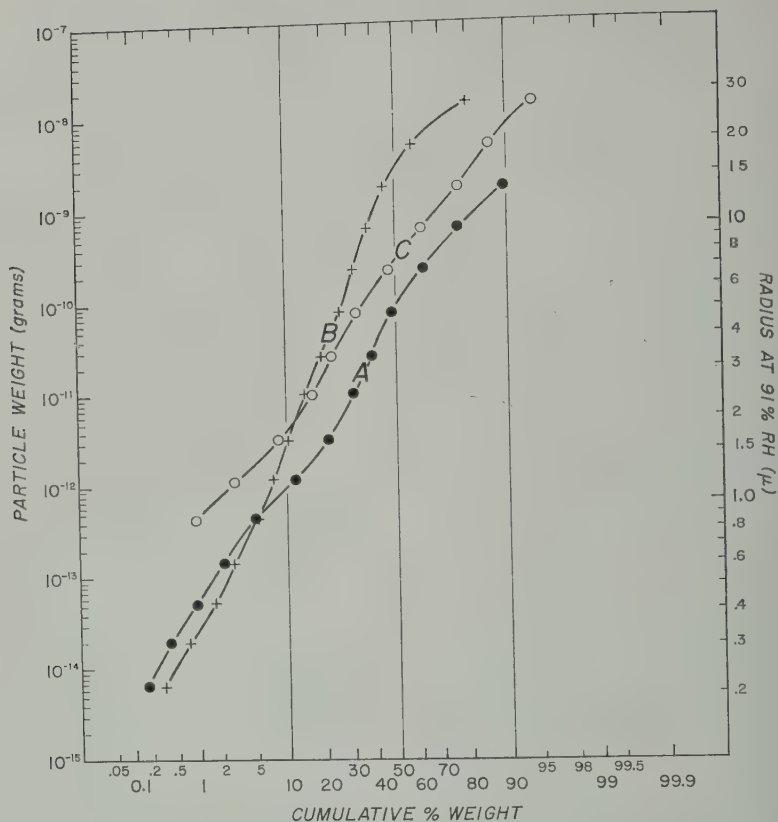


Fig. 9. Cumulative weight per cent distribution among steam-cloud particles formed in laboratory. Curve A represents particles formed when sea water was sprayed on hot lava, and curve B represents those formed when a saturated sea-salt solution was used. C represents distribution among the average volcano steam-cloud particles of Hawaii.

the hot lava enters the sea are shown in plates 1 and 3 of *MacDonald and Hubbard* [1951]. Plate 1 shows a large dense cloud rising when lava, advancing slowly along a broad front, encounters the sea. Plate 3 shows a negligible amount of steam around 'a stream of liquid rock gliding smoothly into the sea ...' In this latter example, the steady flow of fluid is thought to have formed a solidified lava 'tube' under water, which insulates the molten interior and greatly reduces the rate of transfer of heat to the sea as the lava continues its flow down the submarine slope. In contrast, plate 1 (and Fig. 1) is thought to represent a condition in which relatively small pieces or volumes of lava enter the sea intermittently.⁴ In this case it is thought that lava tubes are not formed and that many

small incandescent surfaces are exposed to dissipating much heat in the immediate vaporization of sea water. No doubt the slopes of shore and of the sea bottom influence the details of the mechanics of the encounter of lava with the sea, in addition to the effects of the viscosity of the lava, its gas content, and the nature of the surface which it is arriving at the shore.

Effects of sea state on steam-cloud formation. If the sea is calm, it is thought that the details of the area of the zone of contact between the sea and lava will be narrow. Heavy surf along the shore should act to broaden this zone of contact, and thus to wet hot lava more briefly but repeatedly to wet hot lava with relatively small amounts of water. As its flow finally brings about complete vaporization and emulsion, the heavy seas will increase the vaporization and atomization of water and salt above the calm conditions during calms.

⁴ The presence of the steam makes it difficult if not impossible to see the details of the entry or contact process.

cts of winds on cloud. In addition to the et effects of winds in causing waves on the hey can alter the steam-cloud particle tration in at least two other ways.

t, winds can act to produce turbulent of the clouds with the ambient air and o alter the rates of cooling of the sodium e vapor rising in some of them. This cooling rate will affect the salt crystal time and, as pointed out by *Schaefer* (p. 82), cause differences in the sizes of the s produced.

ds can alter the production of particles in r way—through the transport of spray ge salt particles from the lava-sea contact c back over the approaching hot lava s and fields. This requires a certain wind m, of course. Thus the larger spray and rticles, produced in the zones of explosive and vaporization, would fall out of the clouds onto an even hotter surface, where ould be vaporized and would probably haller airborne particles.

itions resulting in contact between con- ed solutions of sea salts and hot lava can, wn in Figure 8, produce many more s than are formed when the presence of ater rapidly cools the lava surface. It is ed that some of this increase in number sult from the vaporization and subsequent eation of sodium chloride, which has a rable vapor pressure at temperatures of 1000°C.

WEIGHT DISTRIBUTION AMONG THE SALT PARTICLES

relative weight per cent distributions were from the data given in Figure 8, curves B, and Figure 4, curve A. The results of computations are shown in Figure 9. e at the median mass particle of curve A, e represents a sea-water steam cloud, is 4 order of magnitude smaller than the mass particle representing a cloud from ated sea salt solution (curve B). This idicates that the particle size distribution d probably differs from beginning to end sh boiling process as the salt concentra- eases. Curve C of Figure 9, which ts an average of the natural steam-cloud (i.e., curve A, Fig. 4), falls in major een the two curves of the artificially

formed particles. This result is tentatively interpreted to mean that the natural clouds were, at this time, largely a mixture of particles arising from the flash boiling of waters ranging in salt concentration from that of sea water to saturated solutions. The deficiencies of curve C for the smaller particles reflects the failure of the sampling technique used, as discussed.

8. CONCLUDING REMARKS

Meteorologists, geochemists, and others are increasingly concerned with information about salt particles from the sea which become airborne. These particles are thought to be essential in some cloud physics and other atmospheric processes, and in geochemistry, and are known to arise from bubbles present at the surface of the oceans [*Blanchard and Woodcock*, 1957]. It is important to explore more thoroughly the contribution of relatively short-term sources of these particles, such as volcanic activity at or near sea level.

Wexler [1953, p. 95] and *Menzel* [1953, p. 119] have noted the possible role of the freezing nuclei in volcanic dusts in causing large departures from average snowfall in the atmosphere, and in initiating long-term weather changes. The enormous production of giant salt nuclei by the lava-sea contact, and the potential role of these nuclei in the initial formation of raindrops, cause the authors to think that steam-cloud particles should be added to the growing list of factors influencing long-term weather changes.

Concerning the variability of cloud density, it is pertinent to state here that comparison of the range of rate of salt particle production on hot lava and the average of the oceans, such as that made at the end of section 5 above, are of limited meaning at this time. It seems likely that far greater rates will be found when it becomes possible to obtain samples from steam clouds that rise from the porous spongelike aa lava and especially from areas where explosive shattering of the lava surface occurs. A maximum rate of exchange of heat from the lava to the sea water and the air should attend the resulting rapid exposure of multiple hot surfaces. It is possible that a few square kilometers of hot lava, under certain conditions of contact with the sea, can produce a weight of sea-salt nuclei per unit of time equal to a large fraction of the total average production of all of the oceans of the

earth. An occurrence such as this, in a relatively limited area, would be expected to produce changes in the radiation, precipitation, and thermodynamic characteristics of the atmosphere.

The results given in this paper are regarded as a beginning in the study of the salt-aerosol-producing capacity of the earth's lava-sea-atmosphere contact zones. In the future expansion of knowledge of this 'capacity' and its consequences, the following tentative steps seem pertinent at this time:

(a) The chemistry and the sizes of the aerosols produced by vapors arising from dry sea salts at various high temperatures should be investigated. This can be done in the laboratory.

(b) When eruptions and steam clouds occur, particle samples and counts of condensation nuclei should be made in the clouds by aircraft at various short distances from the source areas. The trapping of small volumes of these clouds seems to be the best method of obtaining these samples, since the number and size range of the particles make representative sampling by impingement alone very difficult. A study of the aerosol content, size, and rate of ascent of the steam clouds may be used to determine the variability of particle production rate under the changing conditions of steam-cloud formation by the lava and to test the estimates made in section 5 above.

(c) The great aerosol clouds from the lava-sea-atmosphere contact zones should be followed by specially instrumented aircraft to learn in what way events and properties within the air masses concerned mark them as unusual. It seems reasonable to regard these clouds as natural large-scale salt-seeding experiments, offering extraordinary observational opportunities. Study of the history of weather and other changes associated with these clouds, but far away from the source, may also give us some further ideas about the meteorological roles of the salt particles that are normally produced by the sea.

Acknowledgments. The authors thank Mr. Saul Price, USWB, Honolulu, and Dr. Harry Wexler, USWB, Washington, for their recognition of the rare opportunities for geophysical studies existing during the 1960 Kilauea eruption and for their efforts in aiding the first author during the field work in Hawaii. The first author also wishes to express his appreciation of the help given by Jerry P. Eaton, W.

U. Ault, and D. H. Richter of the Kilauea Volcano Observatory, U.S.G.S., and Jack Pales and John Foss of the Mauna Loa Observatory, USWB. The flying abilities and helpfulness of Mr. W. Stearns of Murray Air, Ltd., of Hilo, Hawaii, in the sampling flights pleasurable as well as efficient. Criticism of the manuscript by Adrian F. Richter was very useful.

The work was supported by the Office of Naval Research, contract NONR-798(00) (NR-082) and the National Science Foundation, grant 100-10000.

REFERENCES

- Arons, A. B., and C. F. Kientzler, Vapor pressure of sea-salt solutions, *Trans. Am. Geophys. Union*, 722-728, 1954.
- Blanchard, D. C., The electrification of the atmosphere by particles from bubbles in the sea, thesis, Massachusetts Institute of Technology, 1961.
- Blanchard, D. C., and A. H. Woodcock, Bubble formation and modification in the sea and its meteorological significance, *Tellus*, 9, 141-157, 1957.
- Dietz, R. S., and M. J. Sheehy, Transpiration of Myojin volcanic explosions by water sound, *Bull. Geol. Soc. Am.*, 65, 944-954, 1954.
- Eaton, J. P., and K. J. Murata, How volcanoes grow, *Science*, 132, 925-938, 1960.
- Eriksson, Erik, The yearly circulation of sulfur and sulfur dioxide, meteorological, geochemical and geological implications, 1, *Tellus*, 11, 375-403, 1959.
- Higashi, K., K. Nakamura, and R. Hara, Thermodynamic gravities and the vapor pressures of concentrated sea water at 0-175°C, *J. Soc. Chem. Japan*, 34 (Abstr. 72B), 166-172, 1931.
- Landsberg, H., Atmospheric condensation in the Gerlands Beir, *Geophys., Suppl. 3*: 155-253, 1954.
- Langmuir, I., and K. B. Blodgett, A mathematical investigation of water droplet trajectories, *Elec. Research Lab. Rept.*, Air Technical Service Command, contract W-33-038-AC-9151, 1945.
- Langmuir, Irving, Super-cooled water droplets and rising currents of cold saturated air, in *Physical Static Studies, Gen. Elec. Research Lab.*, Air Technical Service Command, contract W-33-038-AC-106-SC-65, pp. 1-91, 1944.
- MacDonald, G. A., and D. H. Hubbard, Volcanic activity of Hawaii National Park, *Hawaii Naturalist*, vol. 4, no. 2, 43 pp. Naturalist Division, U.S. National Park, 1951.
- Mason, B. J., *The Physics of Clouds*, Clarendon Press, Oxford, England, 1957.
- Menzel, D. H., On the causes of the ice age, in *Climatic Change. Evidence, Causes and Remedies*, edited by Harlow Shapley, Harvard University Press, Cambridge, Mass., pp.117-121, 1959.
- Schadt, C., and R. D. Cadle, Critical comparison of collection efficiencies of commonly used aerosol sampling devices, *Anal. Chem.*, 29, 864-868, 1957.
- Schaefer, V. J., Laboratory, field and flight

- ts, *Final Report Project Cirrus, 1, R L 785*, General Electric Co., Schenectady, N. Y., 1953.
- on, R. H., A phenomenal haze in the Pacific, *otherwise*, 83-84, 1950.
- s, G. J. (editor), The eruption of Krakatoa subsequent phenomena, *Report of the Krakatoa Committee of the Royal Society*, London, 1888.
- ry, S., The identification of individual hygroscopic particles in the atmosphere by the phase-solubility method, *J. Appl. Phys.*, 24, 1099-1102, 1953.
- ey, S., The composition of hygroscopic particles in the atmosphere, *J. Meteorol.*, 11, 334-338, 1952.
- t, H., On the causes of the ice ages, in *Climatic Change—Evidence, Causes and Effects*, edited by Harlow Shapley, Harvard University Press, Cambridge, Mass., pp. 73-105, 1953.
- Woodcock, A. H., and Mary M. Gifford, Sampling atmospheric sea-salt nuclei over the ocean, *J. Marine Research, Sears Foundation*, 8, 177-197, 1952.
- Woodcock, A. H., Atmospheric salt particles and raindrops, *J. Meteorol.*, 9, 200-212, 1952.
- Woodcock, A. H., and D. C. Blanchard, Tests of the salt-nuclei hypothesis of rain formation, *Tellus*, 7, 437-448, 1955.
- Woodcock, A. H., The origin of trade-wind orographic shower rains, *Tellus*, 12, 315-326, 1960.

(Manuscript received May 8, 1961.)



A Note on the Turbulence Generated by Gravity Waves

O. M. PHILLIPS

*Mechanics Department
The Johns Hopkins University, Baltimore, Maryland¹*

Abstract. When gravity waves move across the surface of a liquid of small viscosity, an angular vorticity (turbulence) field is generated which can attain a statistically steady state through the balance of vorticity generation by the straining associated with the waves and of viscous diffusion. The influence of this vorticity field on the wave motion is examined. It is found that in the equations describing the fluctuating properties of the wave motion, the influence is usually of the third order in the wave slope and results in a slow attenuation of the waves of the form $a(t) = a(t_0) \{1 + \text{const.} (t - t_0)\}^{-1/2}$, where $a(t)$ is the wave amplitude at time t .

Introduction. Although the use of potential theory has been very successful in describing many aspects of the dynamics of gravity waves, it is known that in a real fluid the motion cannot be truly irrotational. Longuet-Higgins [1960] has shown that if the motion starts from rest, the first-order² vorticity is zero except in a boundary layer near the free surface, and possibly at the bottom. Outside the boundary layer, the vorticity is initially zero, but ultimately it settles down to a mean value of second order. This behavior results in, for example, a transport velocity in a two-dimensional wave which is different from the value found by potential motion, even when the fluid viscosity is vanishingly small. In oceanographic applications, one cannot expect the water motion to be exactly two-dimensional, even if the wave is uni-directional. Any stray disturbances result in a distortion of the vortex field, which can be amplified by the straining of fluid elements that is associated with the wave motion.

It appears that in this, as in other situations in physical fluid mechanics, the presence of vorticity is one of the facts of nature, and it

is important to investigate the influence of this random vorticity distribution upon the wave dynamics. The immediate motivation for the present study arises from the current interest in the higher-order, nonlinear effects present in a wave field and the desire to know to what order the assumption of potential motion in the water is consistent with the presence of a weak vortex (turbulent) field that is generated and maintained by the wave motion. This is the aim of this note, and it is found that the turbulent vortex field maintained by the wave motion influences the momentum equation at the third order in the wave slope and results in an attenuation of the wave system.

2. The vorticity field. Consider a vorticity distribution near the surface of a liquid across which is moving a periodic gravity wave. The wave motion convects and distorts the vortex lines, and ultimately a statistically steady state will be attained with the generation of vorticity by the stretching of vortex lines being balanced by the losses through viscous diffusion. In other words, the energy of the turbulent vortex motion is maintained by the balance between energy input from the wave motion and viscous dissipation into heat, and in the statistically steady state the intensity and distribution of the vortex field adjusts itself so that this balance is maintained. The detailed analysis of this situation presents some difficulties, but fortunately for our purpose it suffices to have only an order-of-magnitude estimate of $\overline{\omega^2}$, the mean square vorticity maintained by the wave field. Such an

¹ Present address: St. John's College, Cambridge,

² Basic zero-order quantities are the wave amplitude and the wave frequency, and the surface elevation ξ is of first order. Thus the fluid velocity u and the velocity potential ϕ are of first order, while products of any two first-order quantities are of second order, etc.

estimate is sought in the following investigation.

In the vorticity equation,

$$\frac{d\omega_i}{dt} = \omega_i \frac{\partial q_i}{\partial x_i} + \nu \nabla^2 \omega_i$$

where q_i represents the velocity components of the fluid and ν the molecular viscosity, the first term on the right-hand side represents the rate of generation by the stretching of vortex lines. If we multiply this equation by ω_i and average, we obtain

$$\overline{\frac{d}{dt} \omega^2} = \overline{\omega_i \omega_j \left(\frac{\partial q_i}{\partial x_j} + \frac{\partial q_j}{\partial x_i} \right)} + 2\nu \overline{\omega_i \nabla^2 \omega_i}$$

where the generation is now expressed in terms of the rate of strain

$$e_{ij} = \frac{\partial q_i}{\partial x_j} + \frac{\partial q_j}{\partial x_i}$$

Near the water surface in a field of gravity waves, the rate of strain can be regarded as the sum of three contributions, (i) an oscillatory rate of strain $(e_{ij})_w$ resulting from the wave motion, of order akn ,³ where a , k , and n represent respectively the amplitude, wave number, and frequency of the gravity wave, (ii) a mean rate of strain $(E_{ij})_w$ associated with the wave motion in a fluid of small viscosity, as shown by *Longuet-Higgins* [1953], which is of order $a^2 k^2 n$, and (iii) a random rate of strain $(e_{ij})_T$ associated with the turbulence, and of order $(\overline{\omega^2})^{1/2}$.⁴ The rate of generation of $\overline{\omega^2}$ can thus be expressed as

$$\overline{\omega_i \omega_j (E_{ij})_w} + \overline{\omega_i \omega_j (e_{ij})_w} + \overline{\omega_i \omega_j (e_{ij})_T} \quad (1)$$

Let us first calculate the order of magnitude of the generation of mean square vorticity by the

³ For if the wave profile $\xi \sim a \cos(kx - nt)$, then the velocities near the surface are of order $\dot{\xi} \sim an \sin(kx - nt)$, and the strain rate, proportional to the spatial velocity derivatives, are of order $akn \cos(kx - nt)$.

⁴ The separation of these three can be achieved by consideration of the method of averaging. By fixing our point of observation with respect to the wave profile and taking an ensemble average, the turbulent component becomes the difference between the instantaneous rate of strain and this average value. The oscillating component (i) is given by the variation of this average as the phase of the wave relative to the observation point is varied, and the component (ii) by its mean, i.e. a 'complete average.' The overbars in (1) denote 'complete averages' of this kind.

mean rate of strain, described by the first of these terms. At a depth z below the equilibrium free surface, the rate of generation of $\overline{\omega^2}$ resulting from the mean rate of strain is of order

$$\overline{\omega^2} a^2 k^2 n \exp(kz)$$

This is balanced, in a statistically steady state, by the term $2\nu \overline{\omega_i \nabla^2 \omega_i}$, where ν is the kinematic viscosity of the water. The predominant physical mechanism described by this term is the balance of viscous diffusion and destruction of vorticity, which is of order $\nu \overline{\omega^2} \eta^{-2}$, where η is the Kolmogoroff microscale or, roughly, the scale of the smallest eddies of the turbulence [see, for example, *Batchelor*, 1953]. The gradient transport of vorticity is of order $\nu \overline{\omega^2} k^2$ (since the turbulent field extends over a depth of order k^{-1}) which is smaller than the former contribution by a factor $(\eta k)^2 \ll 1$. The Kolmogoroff microscale $\eta \sim (\nu^3/\epsilon)^{1/4}$, where ϵ is the rate of energy dissipation per unit mass, and in view of the known (exact) relation $\epsilon = -\nu \overline{\omega^2}$ [Batchelor 1953] it follows that $\nu \eta^{-2} \sim \overline{\omega^2}^{1/2}$. We then have

$$\overline{\omega^2} a^2 k^2 n \exp(kz) \sim (\overline{\omega^2})^{3/2}$$

or

$$(\overline{\omega^2})^{1/2} \sim a^2 k^2 n \exp(kz)$$

According to this preliminary estimate it appears that the random vorticity field is of second order in terms of the wave parameters and so is of the same order as the mean vorticity field. We must now examine whether this is consistent with the presence of the other two terms in (1). The first term of (1), the one we have already considered, is near the surface of order

$$\overline{\omega^2} (E_{ij})_w \sim a^6 k^6 n^3$$

i.e. of the sixth order. The last term, in view of the turbulent rate of strain is of order $\overline{\omega^2}^{3/2}$ so from (2) is likewise of sixth order. The turbulence then provides a contribution which is comparable with that from the mean rate of strain but does not affect the order-of-magnitude estimate (2).

The remaining term of (1) involves the covariance between the vorticity field and the cyclic rate of strain $(e_{ij})_w$. Our first reaction is that, since the net strain over a cycle is zero,

of vortex generation from this strain be, if not precisely zero, at least very small. A slightly closer study tends to confirm this suggestion. The maximum strain (over half a wavelength) is of order ak , so that the fractional increase in vorticity along the direction of the strain is also of order ak . If this increase of vorticity is perfectly correlated with the increase of strain, then

$$\overline{\omega_i \omega_j (e_{ij})_w} \sim \overline{\omega^2} (ak) [\overline{(e_{ij})^2}]^{1/2} \\ \sim a^6 k^6 n^3$$

(3). However, this is likely to be an overestimate, since the vorticity increment is more weakly correlated to the strain itself, whose phase differs by $\pi/2$ from that of the strain rate; presently the numerical value of the correlation is probably very low.

It seems, then, that the dominant mechanisms of vortex generation are the stretching of vorticity by the mean rate of strain associated with the wave field and by the turbulence itself, the latter being of the same order of importance. Expression (2) appears to give a plausible first-order estimate of the order of magnitude of the root mean square vorticity field. It is interesting to note that the mean vorticity field is of second order, according to Longuet-Higgins [1953] and is independent of the (small) wave amplitude.

We now study the role played by this vorticity field in the momentum/wave dynamics. The momentum equation can be written in the form

$$\rho \left(\frac{d\mathbf{q}}{dt} + g\mathbf{z} + \frac{1}{2} \nabla q^2 \right) - \mathbf{q} \times \boldsymbol{\omega} \\ = -\nu \nabla \times \boldsymbol{\omega} \quad (3)$$

where \mathbf{q} represents the fluid velocity, and we wish to determine the order of magnitude of the perturbations resulting from the vorticity field. The turbulent velocity disturbance can be estimated most readily by using the empirical

$$\epsilon \approx -\frac{q_i^3}{l} \quad (4)$$

where $[\overline{q_i^2}]^{1/2}$ is the root mean square of the turbulent velocity vector and l is the length scale of the energy-containing eddies. A relation of this type has been found to hold under a wide

variety of conditions; some of the evidence is summarized by Batchelor [1953, §6.1]. Since also $\epsilon = -\nu \overline{\omega^2}$, we have from (4)

$$q_i \sim l^{1/3} \nu^{1/3} (\overline{\omega^2})^{1/3} \quad (5)$$

The scale l of the energy-containing eddies cannot exceed the thickness of the turbulent layer, k^{-1} , and is probably considerably smaller, so that replacing $l^{1/3}$ by $k^{-1/3}$ in (5) will result in rather an overestimate for q_i . Making this substitution, and with the aid of (2), we find that

$$q_i \sim c(ak)^{4/3} R^{-1/3} \quad (6)$$

at most, where $c = n/k$ is the phase velocity of the waves and $R = n/\nu k^2$ is a 'Reynolds number' based on the wavelength and phase velocity of the gravity wave. The basic fluid velocity associated with the wave is of first order, i.e. $c(ak)$, so that the disturbance in the term $\frac{1}{2} \nabla q^2$ is of the order of k times the basic particle velocity, ck times the disturbance velocity (6), or of order $c^2 k (ak)^{7/3} R^{-1/3}$. There is also the second-order mean velocity field [Longuet-Higgins, 1953] that is steady in time with respect to the wave profile. This introduces a perturbation in the term $\frac{1}{2} \nabla q^2$ of third order, i.e. of order $c^2 k (ak)^3$, and the acceleration term, measured in a frame of reference at rest, resulting from this motion can readily be shown to be of third order also. The term on the right of equation (3) is of order $\nu (\overline{\omega^2})^{1/2} \eta^{-1}$ since the vorticity spectrum is concentrated near the Kolmogoroff wave number, η^{-1} , or, using the expressions given previously, of order $c^2 k (ak) R^{-3/2}$.

Table 1 can thus be constructed, and it is evident that the magnitude of the perturbation introduced by the vorticity field depends on the wave 'Reynolds number.' In all oceanographic contexts, R is very large indeed. For a wave of wavelength 1 m, $R \sim 10^6$, so that the terms involving 'Reynolds number' are small compared with the third-order terms if the wave slopes exceed 10^{-3} , in this case. For longer waves, R is even larger.

It seems, then, that under most oceanographic conditions, the largest disturbance terms introduced by a vorticity field in statistical equilibrium with the waves are of order $(ak)^3$, so that

⁵ The quotes are intended to remind the reader that, although R has the structure of a Reynolds number, it does not have the clear dynamical interpretation given in most viscous flow problems.

the perturbation in the quantity $\nabla(p/\rho + gz)$ from the value given by potential theory must also be of third order at most. It follows, therefore, that as far as the time-dependent properties of the wave field are concerned, the presence of the vorticity field will be felt at the third order, and that it is physically unrealistic to carry calculations based on potential theory beyond this order, expecting meaningful results.

An even stronger warning must be sounded in connection with many oceanographical situations. If any wind is blowing over the water surface, we have in addition not only pressure fluctuations generating waves but also the mean and fluctuating shear stresses that generate additional vorticity. In such cases the vortex field may be even larger than that estimated here, but, since its magnitude and distribution depends on a number of additional parameters in a way as yet imperfectly understood, a detailed investigation seems premature.

What is the nature of the influence of a random vorticity field on the wave motion? The length scales of the turbulence near the surface, generated either by the wave field itself or by wind shear stresses or by wave breaking, are comparable with or less than the wavelength of the gravity wave, so that one is at first inclined to expect its influence to be qualitatively similar to that of an 'eddy viscosity.' Dr. Harold Grant of the Pacific Naval Laboratory pointed out in a letter to me that one cannot take this concept too literally.⁶ However, the turbulent vortex field *does* extract energy from the wave motion and dissipates it to heat. (Irregular motions whose length scales are large compared with the wavelength result, of course, in scattering of the gravity wave [Phillips, 1959]). If the vorticity field is maintained by the wave motion, and *not by any wind*, then we have rather an interesting situation like that in turbulent shear flows where the magnitude of the 'eddy viscosity' depends on the mean motion. Here, the dependence is on the amplitude of the wave field itself, and the wave attenuation will not be of the exponential

TABLE 1. Orders of Magnitude of the Perturbations in the Momentum Equation Associated with the Wave-Induced Vorticity Field (All terms are multiplied by the zero-order factor c^2k .)

Term	Order of Irrotational Contribution	Lowest Order of Vorticity Contribution
$\partial \mathbf{q} / \partial t$	ak	$(ak)^3, (ak)^{7/3}, \dots$
$\nabla(\frac{1}{2}q^2)$	$(ak)^2$	$(ak)^3, (ak)^{7/3}, \dots$
$\mathbf{q} \times \boldsymbol{\omega}$		$(ak)^3$
$\nu \nabla \times \boldsymbol{\omega}$		$(ak)R^{-3/2}$

type when the viscosity parameter is considered. An explicit expression for the wave attenuation can be found simply.

In a statistically steady state with no mean flow, the rate of energy loss per unit area from waves is equal to the rate of energy dissipation per unit area by the turbulence, given by

$$\frac{dE}{dt} = - \int_0^\infty \rho \epsilon dz = -\nu \rho \int_0^\infty \bar{\omega}^2 dz \sim -\frac{\nu \rho \omega}{g}$$

from (2). Since the wave energy $E \approx \frac{1}{2} \rho g a^2$ per unit area, it follows that

$$\begin{aligned} \frac{d}{dt} a(t) &\sim -\nu a^2 k^3 n^2 / g \\ &\sim -R^{-1} a^3 k^2 n \end{aligned}$$

where $R = n/\nu k^2$ and $n^2 = gk$ for gravity waves in deep water. Thus

$$a(t) \sim a(t_0) \{1 + R^{-1} a^2(t_0) k^2 n (t - t_0)\}^{-1/2}$$

where $a(t_0)$ is the wave amplitude at time t_0 .

Although this attenuation is the result of third-order disturbances to the momentum equation, the presence of the inverse 'Reynolds number' in (8) implies that this is a slow process compared with some other third-order dependent properties of the wave field. The attenuation time, or the time for the amplitude to be reduced to $1/\sqrt{2}$ of its original value is

$$\theta_a \sim \frac{R}{2\pi} \{a(t_0)k\}^{-2} \text{ wave periods}$$

⁶ He and Dr. R. W. Stewart have also considered in an unpublished work the influence of turbulence on the wave motion, arriving at some conclusions in agreement with those here, but finding an attenuation rate different from (6). His technique extrapolates from theories of the sudden distortion of turbulence to the nearly cyclic straining characteristic of wave motion.

the other hand, the third-order nonlinear interactions between two primary waves of amplitudes α , β , wave numbers \mathbf{k}_0 , \mathbf{k}_1 , and frequencies n_0 , n_1 , can result [Phillips, 1960] in development of a tertiary wave, of wave number $2\mathbf{k}_0 - \mathbf{k}_1$, whose amplitude becomes comparable with those of the primary wave in cycles of the tertiary wave, where

$$\theta_i = (2\pi)^{-1}(\alpha k_0)^{-1}(\beta k_1)^{-1} \quad (10)$$

the amplitudes and wave numbers involved comparable in the two cases, the attenuation is larger than the interaction time by a factor R , which is usually very large. This gives no reason to believe that potential wave theory, if carried to the third order, will give numerical results not only of qualitative but of quantitative interest, even perhaps when

the vorticity field is rather more intense than that discussed here, and is maintained partly by wind stresses.

REFERENCES

- Batchelor, G. K., *The Theory of Homogeneous Turbulence*, Cambridge University Press, 1953.
 Longuet-Higgins, M. S., Mass transport in water waves, *Phil. Trans. Roy. Soc. London A*, **245**, 535-581, 1953.
 Longuet-Higgins, M. S., Mass transport in the boundary layer at a free oscillating surface, *J. Fluid Mech.*, **8**, 293-306, 1960.
 Phillips, O. M., The scattering of gravity waves by turbulence, *J. Fluid Mech.*, **5**, 177-192, 1959.
 Phillips, O. M., On the dynamics of unsteady gravity waves of finite amplitude, *J. Fluid Mech.*, **9**, 193-217, 1960.

(Manuscript received December 8, 1960;
 revised June 13, 1961.)

Long-Period Surface Waves from the Chilean Earthquake of May 22, 1960, Recorded on Linear Strain Seismographs¹

JAMES N. BRUNE,² HUGO BENIOFF,³ AND MAURICE EWING²

*Lamont Geological Observatory, Columbia University
Palisades, New York*
and

*Division of Geological Sciences, California Institute of Technology
Pasadena, California*

Abstract. Phase and group velocities of mantle Love and Rayleigh waves obtained from strain seismograph records of the Chilean earthquake are presented. The velocities of mantle Rayleigh waves of period from 300 to 550 seconds agree with those predicted from periods of free spheroidal oscillation of the earth and do not show a flattening of the group velocity curve for periods greater than 380 seconds. Group velocities for mantle Rayleigh waves reach a maximum of 7.8 km/sec at a period of about 1000 sec. Study of initial phases of Rayleigh waves indicates a difference of phase π between the azimuth to Isabella and the azimuths to Ñaña and Ogdensburg. Determinations of phase and group velocities of Love waves have been extended to periods of 700 seconds. The phase velocity data of Satô [1958] has been corrected for the polar phase shift. The correct curve has been identified from the numerous possible curves which result from a 2π ambiguity in the phase correlation made by Satô. Values of phase velocities are presented for periods in the range of 300 to 700 seconds. The group and phase velocities of both Love waves and Rayleigh waves agree well with those predicted for the Gutenberg-Bullen A model of the earth. It is verified that analysis of seismograms in terms of progressive wave trains is equivalent to analysis in terms of standing waves. In the presence of absorption, as for the earth, the analysis in terms of progressive wave trains has many advantages.

Introduction. The studies of long-period Love and Rayleigh waves presented in this paper are based on data obtained by reading directly from seismograms the times of arrival of peaks and troughs of conspicuous wave trains. The seismograms were available from seismometers in Peru, Isabella, California, and Ogdensburg, New Jersey. At Isabella and Ogdensburg the earthquake is recorded directly using displacement seismometers, as has been described by Benioff; at Isabella filters were used to cut out the shorter-period surface waves on one hand and the earth tremors on the other. These seismograms made it possible to study the dispersion of surface waves of periods considerably longer than surface waves previously studied in this way. The results extend well into the range which has hitherto been studied only in terms of standing waves (free vibrations). Phase velocities were determined by correlation of wave trains, over

complete great-circle paths, by a modification of the method described by Nafe and Brune [1960]. The use of phase velocity as the basic information read from the seismogram is preferable to the use of group velocity for two important reasons: First, phase velocity may be determined much more accurately than group velocity by direct measurement of arrival times of peaks and troughs; second, the interpretation of phase velocity data in terms of earth structure is more straightforward than the interpretation of group velocity data.

Once phase velocities and attenuation are known, the dispersed surface wave trains of various orders may be synthesized for any later or earlier time from a Fourier analysis of one even-order and one odd-order train. Initial motion may be determined by making such calculations for the origin time of the earthquake. Information about the initial phases is obtained for the Rayleigh wave trains observed on these records by using the phase velocities presented.

Data from the Isabella, Ñaña, and Ogdensburg strain seismographs. In Figure 1, the records of

¹Lamont Geological Observatory Contribution No. 1028.
²Lamont Geological Observatory.
³California Institute of Technology.

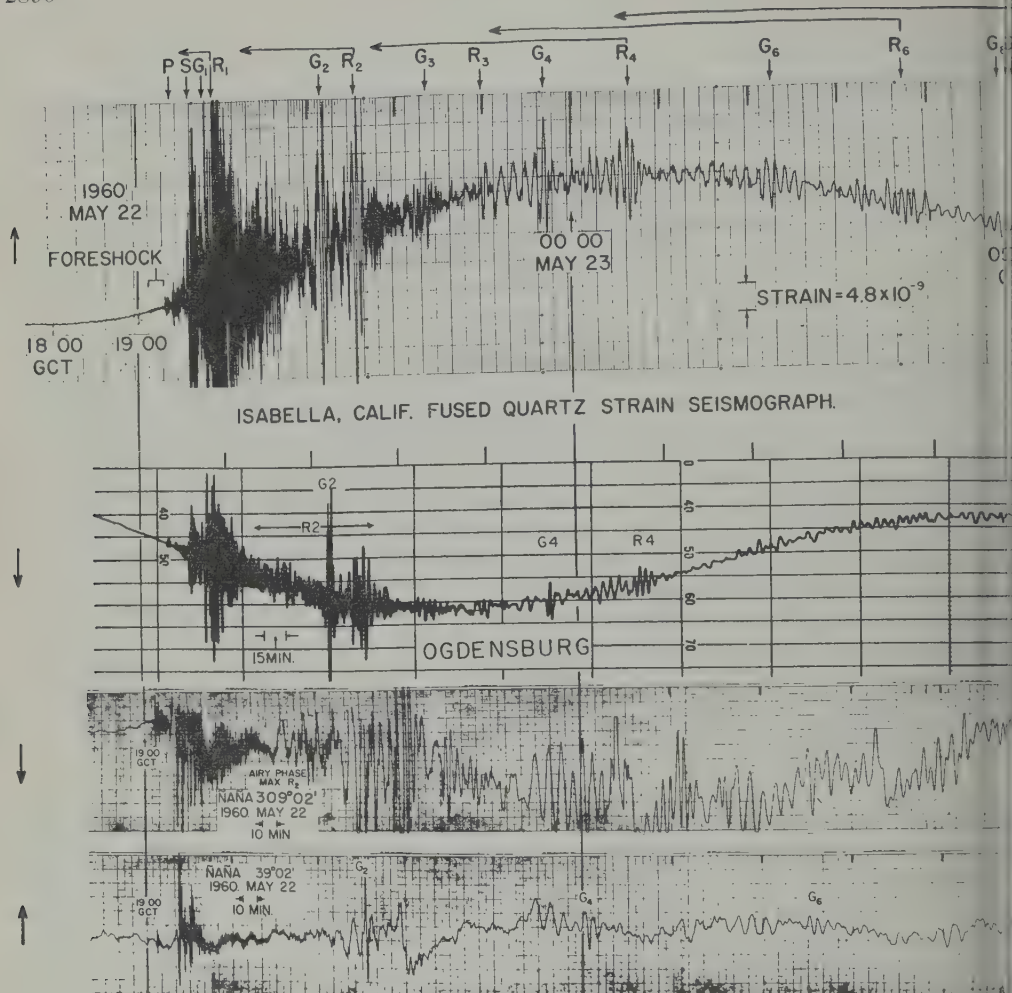


Fig. 1. Seismograms of the great Chilean earthquake of May 22, 1960, recorded with Benioff linear strain seismographs. Table 1 gives epicentral and station data. The Ogdensburg and Isabella seismograms record strain directly, whereas the Ñaña seismograms are filtered, with corner periods of 200 and 500 seconds. 1900, May 22, 0000, May 23, and 0500, May 23, on each record are connected by a solid line.

the Chilean earthquake of May 22, 1960 (see Table 1), from the Benioff linear strain seismographs at Isabella, Ogdensburg, and Ñaña are shown with a common time scale. The direction representing compression for each seismogram is marked by an arrow. The distances from the epicenter (9220, 8744, and 2890 km, respectively) are comparable, facilitating comparison of phases recorded at the three stations. The directions from the epicenter toward the three seismograph stations are indicated in Figure 2, as is the probable trend and extent of the fault, as

deduced by Benioff from the distribution of aftershocks. The direction and approach of surface waves from Chile to Ñaña and orientation of the quartz rod at Ñaña are shown in this figure. The inset indicates the direction of approach of surface waves from Chile compared with the direction of the quartz rod at Ogdensburg and Isabella, respectively. The Ogdensburg instrument is oriented with an azimuth of 30° and is approximately equally sensitive to Love and Rayleigh waves from an azimuth to Chile 180°. The Isabella is

TABLE 1. Station and Epicenter Data
 Epicenter 38°S, 73½°W
 Origin time 19h 11m 20s UT, May 22, 1960
 Magnitude 8½ to 8½

Station	Strike of Quartz Rod	Δ , km	Az. to Chile	Az. from Chile
Isabella, Calif.	328°	9220	146°	306°
Ogdensburg, N. J.	30°	8744	180°	00°
Ñaña, Peru	309°02'	2890	174°	354°
	39°02'			

is oriented along an azimuth of 328° and is much more sensitive to Rayleigh waves than to Love waves from Chile (azimuth to 146°). The two strain seismographs at Isabella are mounted at azimuths 309°02' and 39°02'. The Ñaña records in Figure 1 are filtered with low-pass periods of 200 and 500 seconds. According to the directivity pattern shown in Figure 2 for Isabella, both components are equally sensitive to Love and Rayleigh waves from Chile. Benioff (1950) has noticed previously that the northeast component seems consistently lower in response than the other two to waves coming in a direction equally sensitive to both, even though the calibrations of the three components reveal no difference in sensitivity. At Isabella, the northeast component at Ñaña responds almost exclusively to Love waves, as may be seen from Figure 1.

The first clear event on these seismograms (Figure 1) is the *S* wave from a foreshock (0 equals 15-57). The *P* wave from this foreshock is detectable on the Isabella and the Ñaña seismograms only. Each Love and Rayleigh wave train is marked at, or near, the part of the maximum amplitude for that train. It has long been recognized, the long-period Love waves arrive in short, almost pulslike trains, indicating a group velocity curve which is flat at about 4.4 km/sec over a period range of about 100 to 400 seconds. On the Ñaña seismogram in Figure 1, a train of waves with periods of more than 400 seconds precedes the pulslike train of *G*₂, indicating that the group velocity begins to rise more steeply for these waves. The decaying exponential displacement on the Ñaña 39°02' seismogram at about 22h 00m May 22, just after the *G* wave, is approximately the response of the system to a step

function of strain and is possibly caused by instrumental noise or a failure of the earth near Ñaña upon the arrival of the high-amplitude *R*₂ Airy phase minimum.

It has recently been shown, both by analysis of propagating waves and of standing waves [Brune, Ewing, and Kuo, 1961] that the Rayleigh wave group velocities rise from a minimum of 3.55 km/sec at 225 sec to a maximum of 7.8 km/sec at a period of about 1000 sec. From this very high velocity it is expected that Rayleigh waves of long period could arrive far ahead of the *G* wave of the same order, and even well ahead of *S*. This range of group velocities for *R*, a range of more than 2 to 1, causes each Rayleigh wave train to be very long. The result is overlapping of the trains of different order, which has led to some erroneous identifications in the past. The range of each Rayleigh train in Figure 1 has been indicated by an arrow parallel to the time axis, extending from the arrival time of the minimum of group velocity to that of the maximum. The extent of the overlapping of the *R* trains, even for this earthquake in which the trains of odd order are practically absent, is a full explanation of the complicated character of the Rayleigh waves beyond *R*₂ in typical seismograms from large earthquakes. The long-period waves, which appear ahead of *S* on the Isabella and Ñaña seismograms, have a velocity comparable to the maximum of 7.8 km/sec and are direct evidence of these high group velocities. On the Ogdensburg record the *R*₁ train is weak, but the long *R*₂ train is in excellent agreement with the new dispersion curve [Brune, Ewing, and Kuo, 1961] and has the pulslike *G*₂ wave superimposed just ahead of the part of the Rayleigh wave train with period less than 400 sec. The Rayleigh wave maximum has an impulsive beginning at the time corresponding to a velocity of 7.8 km/sec for *R*₂ on the Ñaña seismogram (for the quartz rod of azimuth 309°02') and a less impulsive but quite clear beginning at the same velocity on the Ogdensburg seismogram. Waves of period near 500 sec appear on the Isabella seismograms just preceding the *G* waves in the orders *R*₂, *R*₄, and *R*₆, also in agreement with the new group velocity data.

The asymmetrical radiation pattern, which makes the even orders of both *R* and *G* much stronger than the odd orders at Ogdensburg as

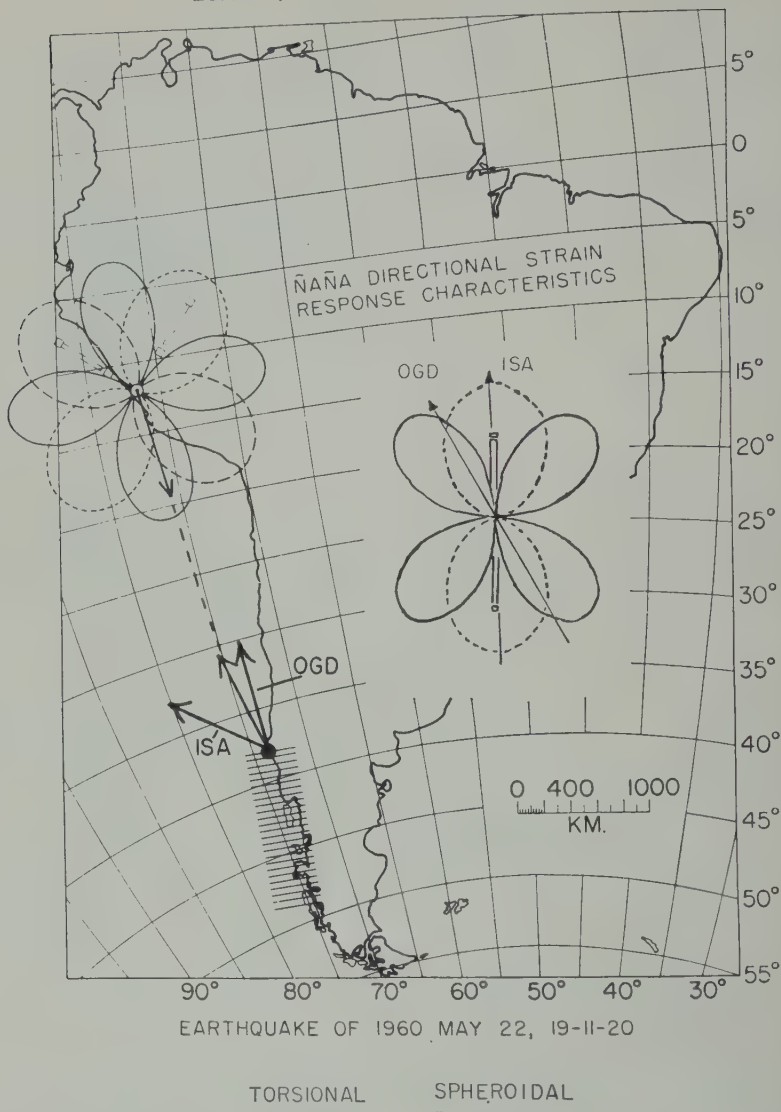


Fig. 2. Direction to strain seismographs from the epicenter of the Chilean earthquake. The crosshatching represents the approximate area covered by aftershocks. The rosettes show the sensitivity of the seismographs to motion of Love and Rayleigh waves as a function of azimuth of approach. The direction of approach from Chile relative to the quartz rod is shown in each case.

well as at Isabella and Ñaña, has been attributed to the effect of a moving source [Benioff, Press, and Smith, 1961]. The spectrum of radiation of the even-order trains is markedly different at Ogdensburg and Isabella. For instance, at Ogdensburg the train R_2 shows high amplitudes for periods near 700 sec, low amplitudes for periods near 500 sec, and high amplitudes again from 300 to 100 sec. In contrast, on the Isabella record, the R_2 train shows high amplitudes for

periods near 500 sec (just preceding the G wave). The absence of these 500-sec waves on the Ogdensburg record is striking.

Analysis of Rayleigh wave data. Rayleigh wave phase and group velocities for theoretical models of the earth have been given by Pekeris, Alterman, and Jarosch [1961], and experimental phase and group velocities derived from observations of free periods of oscillation on the seismographs from the Chilean earthquake have been given

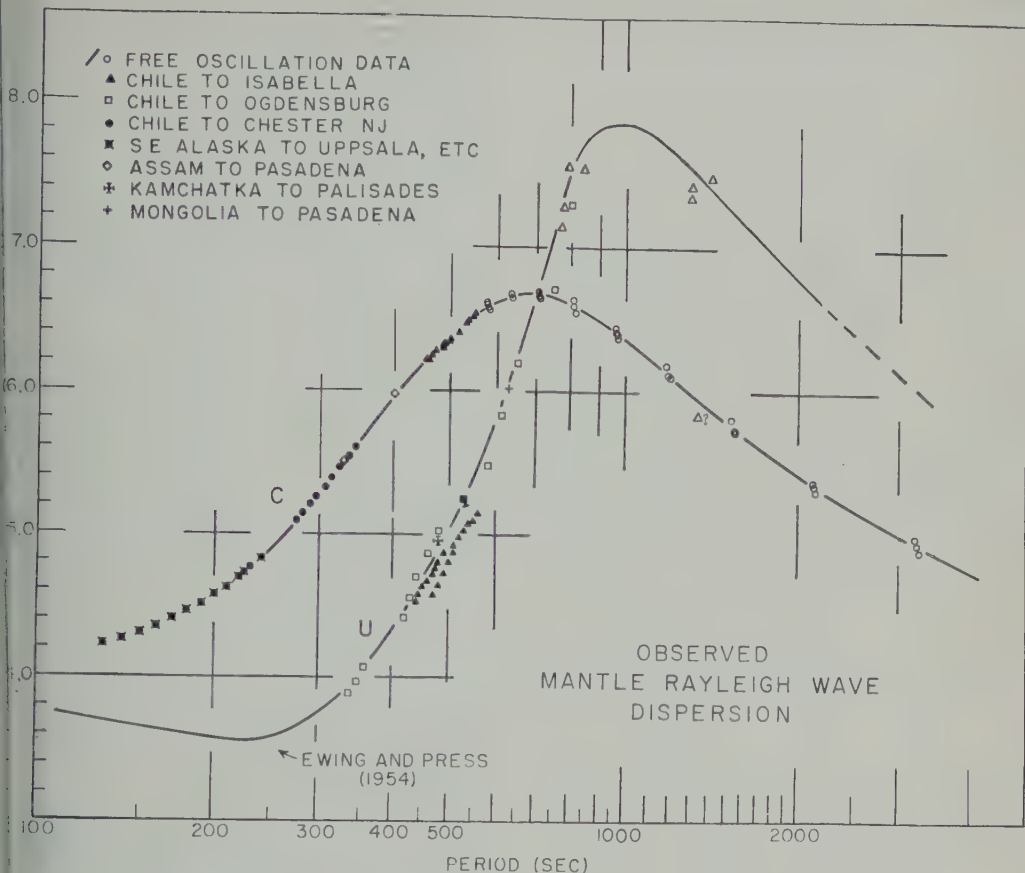


Fig. 3. Phase and group velocity data for mantle Rayleigh waves.

Brune, Ewing, and Kuo [1961]. Periods of spheroidal oscillation of the earth have been determined by frequency analysis of the following records: Los Angeles gravity meter records [Ness, Harrison, and Slichter, 1961]; the records from the Isabella [Benioff, Press, and others, 1961] and Ogdensburg [Alsop, Sutton, and others, 1961] strain seismographs; and the records from Chester [Bogert, 1961], Palisades [Sutton, and Ewing, 1961], and Pasadena [Benioff, Press, and Smith, 1961] long-period seismographs. The theoretical free periods of spheroidal and torsional oscillations are used to identify peaks of the spectra [Brune, Jarosch, and Pekeris, 1961; Pekeris, Brune, and Jarosch, 1961; Bolt and Dorman, 1961; Gilbert and MacDonald, 1960; Satô, Brune, and Ewing, 1960]. The theoretical observed free periods show good agreement, and it is possible to choose between the various

earth models on the basis of the observed free periods. The Gutenberg-Bullen A model [Bolt and Dorman, 1961; Pekeris, Alterman, and Jarosch, 1961] gives better agreement with the observed periods than any other model studied to date.

Phase and group velocities may be obtained from this data by the formula of Jeans [1923], $C = 2\pi a / (n + \frac{1}{2})T$, where C is the phase velocity, a is the radius of the earth, n is the mode number of the free oscillation, and T is the period of the oscillation. Wavelength, λ , is given by $2\pi a / (n + \frac{1}{2})$. Group velocity, U , may be obtained from the relation $U = C - \lambda dC/d\lambda$. It has been shown that the analysis of seismograms in terms of progressive waves may be made consistent with Jean's formula provided a phase advance of $\pi/2$ at each polar or antipodal crossing is taken into account [Brune, Nafe, and Alsop, 1961]. Phase and group velocity curves

TABLE 2. Rayleigh Wave Data from the Isabella Seismogram of the Chilean Earthquake of May 22, 1960

Order	Phase	Time		T , sec	t_1 , sec	U , km/sec
		h	m			
R_2	p	20	50.7	490*	5,964	5.16
	t		54.5	490	6,192	4.97
	p		58.9	490*	6,456	4.77
R_3	p	22	08.2	450*	10,614	4.64
R_4	p	23	03.7	540*	13,944	5.08
	t		08.2	520*	14,214	4.98
	p		12.7	510	14,484	4.89
	t		17.1	500	14,736	4.81
	p		20.9	490	14,976	4.73
	t		25.0	480	15,222	4.65
	p		29.0	470*	15,462	4.58
R_6	p	01	08.7	560*	21,444	5.17
	t		13.3	550	21,720	5.11
	p		17.7	530	21,984	5.04
	t		22.2	520	22,254	4.98
	p		26.7	510	22,524	4.92
	t		30.9	490	22,776	4.87
	p		35.2	480	23,034	4.81
	t		38.7	475	23,244	4.77
	p		42.4	470	23,466	4.72
	t		46.7	460	23,724	4.67
	p		50.8	450	23,970	4.62
	t		54.4	445	24,186	4.58
	p		58.0	440*	24,402	4.54
R_7	p	00	01.0	790*	17,382	7.44
	t		07.6	780	17,778	7.27
	p		14.2	770*	18,174	7.11
R_8	p	00	34.7	840*	19,404	7.77
	t		43.7	840	19,944	7.56
	p		50.2	840*	20,334	7.42
R_{10}	t	02	11.3	1200*	25,200	7.57*
	p		23.2	1290	25,911	7.37
	t		33.2	1380*	26,514	7.20*
R_{14}	p	05	15.7	1400*	36,264	7.47*

* Period poorly determined.

derived from observed spheroidal free periods of oscillation are shown in Figure 3. Calculations based on flat-layer models to fit the group velocity data of *Ewing and Press* [1954a, b] showed that a low-velocity channel in the mantle was necessary to explain the dispersion of mantle Rayleigh waves [*Dorman, Ewing, and Oliver*, 1960; *Takeuchi, Press, and Kobayashi*, 1959]. Observations of *Ewing and Press* [1945b] and

Benioff and Press [1958], which were interpreted as indicating a flattening of the group velocity curve for periods greater than 380 sec, have been revised and are in good agreement with velocities determined from free oscillations [*Brune, Ewing, and Kuo*, 1961]. The long-period Rayleigh wave train R_2 observed on the Ogdenburg strain seismograph gave group velocities in the period range 200 to 800 seconds which agree with the group velocity curve derived from free periods of oscillation.

Phase velocity measurements for many Rayleigh waves from the Assam earthquake were given by *Nafe and Brune* [1960] and revised to take into account the polar phase advances by *Brune, Nafe, and Alsop* [1961]. Phase velocity measurements from the Chilean earthquake recorded at Chester, New Jersey, were given by *Brune, Nafe, and Alsop* [1961]. The measurements were made from a seismogram provided by B. Bogert of the Bell Telephone Laboratories, Murray Hill, New Jersey. These observations of phase velocity are plotted in Figure 3 and are in good agreement with phase velocities determined from the free periods of spheroidal oscillation of the earth.

The data and computations from the record shown in Figure 1 are given in Tables 1 through 4. Table 1 gives the origin time, location, magnitude of the Chilean earthquake and the azimuths and distances to the three recording stations.

Table 2 gives the Rayleigh wave data from the Isabella record shown in Figure 1. The group velocity data are plotted in Figure 3 for the clear trains of period 450 to 500 sec on the Isabella record, R_2 , R_4 , and R_6 . The velocities are consistently slightly below the group velocity curves derived from the above periods of oscillation. This is possibly due to a variation of an initial phase as a function of period. Other long-period data, with periods near 840, 1300, and 1400 sec are subject to large error in measurement of period because one cycle or two of motion appears, but they agree fairly well with the group velocity curves derived from the free oscillations. These long period waves have not been checked for phase motion and are subject to revision when a detailed study is made. Bandpass filtering of the record would make it possible to observe individual wave trains more clearly.

TABLE 3. Computation of Rayleigh Wave Phase Velocities, Chile to Isabella, Complete Great-Circle Path

R_3		Phase	R_4		T , sec	t , sec	$(n - \frac{1}{2})$	$(n - \frac{1}{2})T$, sec	C , km/sec
T , sec	t_1 , sec		T , sec	t_1 , sec					
560	21,444	<i>p</i>	540	13,944	550	7,500	2.5	1375	6.534
550	21,720	<i>p</i>	540	13,944	545	7,776	3	1635	6.517
530	21,984	<i>p</i>	540	13,944	535	8,040	3.5	1873	6.489
520	22,254	<i>p</i>	540	13,944	530	8,310	4	2120	6.465
510	22,524	<i>t</i>	520	14,214	515	8,310	4	2060	6.403
490	22,776	<i>p</i>	510	14,484	500	8,292	4	2000	6.360
490	22,776	<i>t</i>	500	14,736	495	8,040	3.5	1733	6.345
480	23,034	<i>p</i>	490	14,976	485	8,058	3.5	1698	6.292
475	23,244	<i>t</i>	480	15,222	477	8,022	3.5	1670	6.300
470	23,466	<i>t</i>	480	15,222	475	8,244	4	1900	6.308
460	23,724	<i>p</i>	470	15,462	465	8,262	4	1860	6.251
450	23,970	<i>p</i>	470	15,462	460	8,508	4.5	2070	6.216
445	24,186	<i>p</i>	470	15,462	457	8,724	5	2285	6.215
440	24,402	<i>p</i>	470	15,462	455	8,940	5.5	2503	6.217
R_4			R_2						
500	14,736	<i>p</i>	490*	5,964	495	8,772	5	2475	6.355
490	14,976	<i>t</i>	490	6,192	490	8,784	5	2450	6.318
480	15,222	<i>p</i>	490*	6,456	485	8,766	5	2425	6.311
R_{14}			R_{10}						
1380*	36,264	<i>p</i>	1290	25,914	1340	10,350	-2.5	3350	5.842

*Period poorly determined.

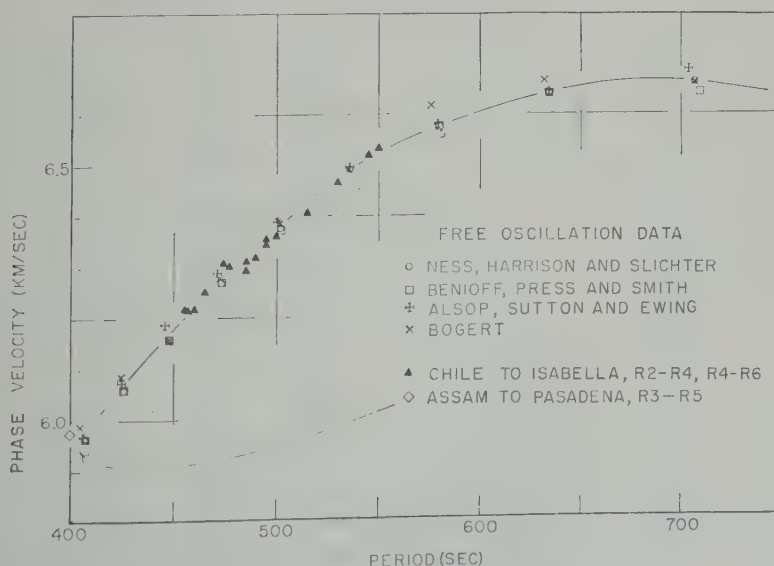


Fig. 4. Detail of Figure 3 showing Rayleigh wave phase velocity data from the Isabella seismogram compared with that obtained from periods of free spheroidal oscillations.

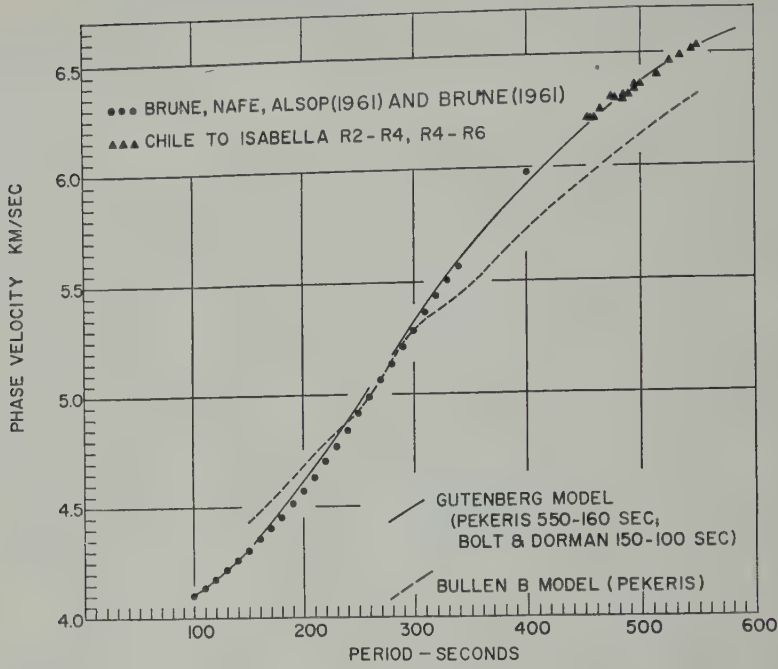


Fig. 5. Rayleigh wave phase velocities compared with theoretical dispersion curves based on the Gutenberg and Jeffreys-Bullen models of the earth.

TABLE 4. Determination of Initial Phase, Chile to Isabella

Order	Phase	T , sec	t_1 , sec	$T/8$, sec	t_2 , sec	C , km/sec	Δ/C , sec	$\Delta/C - t_1$, sec	$\Delta/C - t_2/T$	$\varphi_a/2\pi$	φ_0
R_2	p	520?	5,940	-65	5,875	6.436	4,786	-1089	-2.09	+.91	
	t	500	6,186	-63	6,123	6.364	4,840	-1283	-2.57	+.93	
	p	490	6,456	-61	6,395	6.330	4,866	-1529	-3.12	+.87	
		480?	6,690	-60	6,630	6.296	4,892	-1738	-3.62	+.87	
	p	420	7,674	-53	7,621	6.060	5,083	-2538	-6.04	+.96	
	t	370	7,866	-46	7,820	5.760	5,347	-2473	-6.68	+.82	
	p	320	7,998	-40	7,958	5.438	5,664	-2294	-7.17	+.83	
	t	300	8,154	-38	8,116	5.284	5,829	-2287	-7.62	+.88	
	p	280	8,304	-35	8,269	5.137	5,996	-2273	-8.12	+.88	
	t	240	8,424	*	8,424	4.840	6,364	-2060	-8.59	+.91	
	p	220	8,544	*	8,544	4.702	6,550	-1994	-9.06	+.94	
	t	215	8,652	*	8,652	4.667	6,600	-2052	-9.54	+.96	
	p	215	8,754	*	8,754	4.667	6,600	-2154	-10.02	+.98	
R_3	p	450?	10,614	-56	10,558	6.190	7,955	-2603	-5.78	+.22	+
R_4	p	540	13,944	-68	13,876	6.500	10,895	-2981	-5.52	+.48	+
	t	520	14,214	-65	14,149	6.436	11,004	-3145	-6.05	+.45	+
	p	510	14,484	-64	14,420	6.402	11,062	-3358	-6.58	+.42	+
	t	500	14,736	-63	14,673	6.368	11,121	-3552	-7.10	+.40	+
	p	490	14,976	-61	14,915	6.330	11,185	-3730	-7.61	+.39	+
	t	480	15,222	-60	15,162	6.298	11,245	-3917	-8.16	+.34	+
	p	470	15,462	-59	15,403	6.260	11,313	-4090	-8.70	+.30	+

* This phase was assumed to be part of the airy phase.

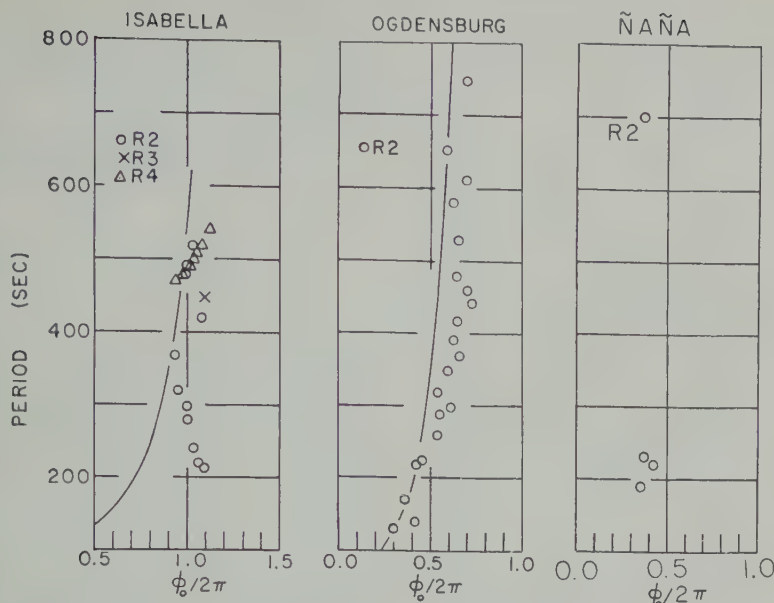


Fig. 6. Initial phases of vertical displacement for Rayleigh waves from the Chilean earthquake. The solid line in the first two sections indicates the predicted variation of initial phase as a function of period for a fault model of the type proposed by Benioff, Press, and Smith [1961].

Table 3 shows calculations of phase velocity from the data in Table 2. Peaks and troughs in the strain trains are correlated (R_2 to R_4), and phase velocity is computed from the formula

$$\Delta/[t - (n - 1/2)T], \quad n = \frac{\varphi_b - \varphi_a}{2\pi} + N$$

where t is the time difference between the two trains, Δ is the distance, φ_a is the phase of the first train and φ_b the phase of the second train, n is an integer which must be determined, and the term $-\frac{1}{2}T$ arises from the polar phase shift of $\pi/2$ at each polar or antipodal passage. In phase velocity determination by this method, the phase shift of the instrument and the initial phase of the origin are eliminated. It also turns out that error in measurement of period does not critically affect the accuracy of the phase velocity measurement, since at the time when a peak of energy of a particular period arrives on the record the phase of the Fourier components contributing to the group varies slowly as a function of period [Brune, 1961]. The integer n is determined from data from more closely spaced stations or from comparisons with theoretical calculations.

From records of the Isabella strain seismograph clear trains of Rayleigh waves in the period

range 450 to 550 sec just preceding G_2 , G_4 , and G_6 are correlated to obtain phase velocities over the great circle passing through the epicenter and Isabella. The phase velocity data plotted in Figure 3 as triangles agree well with the phase velocities determined from periods of free oscillations. This extends the measurement of phase velocity by progressive wave analysis to a period of 550 sec. Figure 4 shows a more detailed graph of this data. Points determined from free oscillation periods are also shown and it is seen that the precision with which phase velocity is determined is approximately the same for the two methods. Precision for both methods can probably be improved considerably with filtering and better techniques of digital analysis. The Rayleigh wave phase velocities available to date are compared with theoretical curves in Figure 5. The comparison of these data with theory in the period range 100 to 350 sec was given by Alsop, Sutton and Ewing [1961] and Bolt and Dorman [1961], who found that the theoretical model of the earth which best fitted the data was the Gutenberg velocity and Bullen A density model. From Figure 5 it is seen that a similar conclusion may be drawn for longer periods, with even better agreement.

The extension of phase velocity measurements

TABLE 5A. Love Wave Data from the Ñaña 39°02' Seismogram of May 22, 1960

Order	Phase	Time		T, sec	t _i , sec	U, km/sec
		h	m			
G ₂	p	21	13.5	720	7,332	5.06
	t		19.0	624	7,662	4.846
	p		24.0	564	7,962	4.663
	t		28.2	510	8,114	4.576
				390		4.389
	p		32.3	270	8,460	4.389
	t		33.6	132	8,538	4.349
	p		34.6	96	8,598	4.318
	t		35.1	66	8,628	4.303
G ₄	p	23	28.2	690	15,414	5.01
	t		33.8	648	15,750	4.90
	p		38.0	600	16,002	4.82
	t		43.1	540	16,308	4.73
	p		46.8	480	16,530	4.67
	t		51.2	450	16,794	4.59
	p		54.9	390	17,016	4.53
	t		58.2	360	17,214	4.48
	p	24	01.1	330	17,388	4.44
	t		04.0	300	17,562	4.39
G ₆	t	01	38.7	690	23,244	5.04
	p		43.9	630	23,556	4.97
	t		48.0	570	23,802	4.92
	p		52.5	540	24,072	4.87
	t		57.2	525	24,354	4.81
	p	02	02.2	510	24,654	4.75
	t		06.2	510	24,894	4.71
	t		27.7	420	26,184	4.48
	p		31.2	360	26,394	4.44
G ₈	t		34.1	348	26,568	4.41
	p		36.5	330	26,712	4.39
	t		39.8	330	26,916	4.35
	p	03	57.3	690	31,560	4.98
	t	04	02.6	660	31,878	4.93
	p		07.8	600	32,190	4.88

of short-period interference, the correlation open to question. If the shorter periods filtered from the record, the correlation would be made more reliable.

Initial phase of Rayleigh waves. Initial phase ϕ_0 is determined from the following relation

$$\varphi_0 = 2\pi(X/CT - t/T + \varphi_s/2\pi + N)$$

$$-\pi/4 - m\pi/2 = \varphi_s - \pi/4 - m\pi$$

where X is the distance, t is the travel time, C is the phase velocity, T is the period, ϕ_0 is the phase of a Fourier component (for non-dispersion $T/8$ sec less than the time of a peak on the record), ϕ_s is the phase of the Fourier component, and the phase $-\pi/4 - m\pi/2$ from the asymptotic expansion of the Legendre polynomial, m being the number of polar antipodal crossings which the wave train has made. The calculations of initial phase for the Rayleigh trains R_2 , R_4 , and R_6 on the Isabella record are shown in Table 4. Similar calculations for the Rayleigh trains R_2 on the Ogdensburg and Ñaña records are filed with the American Documentation Service.

TABLE 5B. Love Wave Data from the Isabella Seismogram of May 22, 1960

Order	Phase	Time		T, sec	U, km/sec
		h	m		
G ₂	t	21	08.6	272	7.7
	p		10.4	222	7.7
	t		12.0	198	7.7
	p		13.3	150	7.7
G ₄	p	23	35.6	366	15.4
	t		38.9	330	16.1
	p		41.5	282	16.1
	t		43.7	198	16.1
G ₆	p	02	11.8	390	25.4
	t		15.0	366	25.4
	p		17.4	360	25.4
	t		20.9	354	25.4
	p		23.9	342	25.4
	t		26.7	312	26.7
G ₈	p	04	47.4	402	34.4
	t		50.5	396	34.4
	p		53.6	390	34.4
	t		57.0	378	35.0
	p	05	00.0	357	35.0
	t		03.2	336	35.0
	p		06.0	318	35.0
	t		08.1	288	35.0

to these long periods by correlation of progressive wave trains emphasizes the agreement of this method with the method of free periods. If wave trains of still longer period can be observed and studied near the epicenter and the antipode, the nonsinusoidal character of spherical harmonic traveling waves should become evident, giving apparent phase velocities which are faster across polar regions than elsewhere [Brune, Wafe, and Alsop, 1961].

A single correlation of a wave of period 1340 sec is tentatively given in Table 3 and plotted with question mark in Figure 3. Because of uncertainty in identification and a large amount

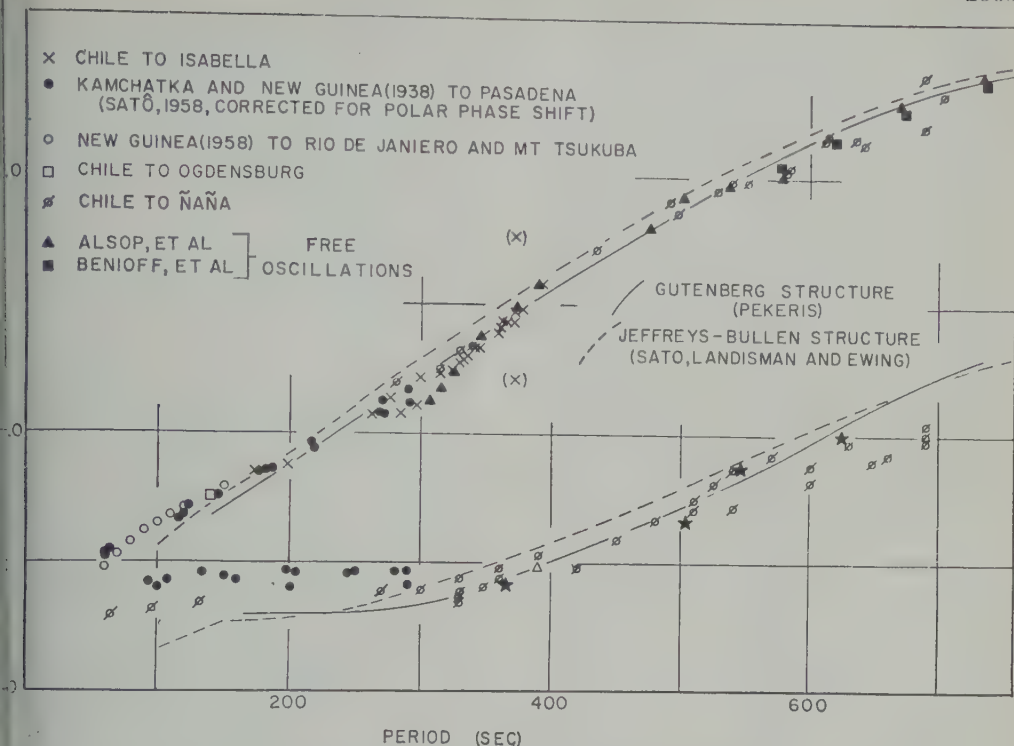


Fig. 7. Love wave phase and group velocity data. Theoretical curves are shown for the Gutenberg and Jeffreys-Bullen (Bullen B) structures of the earth.

ute.⁴ The results are plotted in Figure 6. curve in each figure indicates the variation initial phase as a function of period which l be expected for a source mechanism of ype proposed by Benioff, Press, and Smith for this earthquake, that is, a source ng southward with a velocity of 3 km/sec distance of 1200 km. The variation of phase is calculated from the formula of Menahem [1961]

$$\varphi' = -\frac{\pi b}{CT} (C/v - \cos \theta_0)$$

φ' is the phase shift introduced by the

aterial supplementary to this article has been ed with the ADI Auxiliary Publications t, Photoduplication Service, Library of Con- Washington 25, D. C. A copy may be secured ng the document number 6816 and remitting or 35-mm microfilm. Advance payment is ed. Make check or money order payable to: Photoduplication Service, Library of Con-

propagation of the source, b is the length of propagation (1200 km), v is the rupture velocity (3 km/sec), and θ_0 is the azimuth from the earthquake to the station measured from the direction of propagation.

Very little variation of initial phase as a function of period is predicted at Ogdensburg, and the data are not precise enough to establish the small variation definitely, even though the data show a tendency to vary in the appropriate direction. A somewhat greater variation as a function of period is predicted for Isabella, and in this case the observed initial phase shows no tendency to follow the theoretical curve. Within the limits of error of the measurement, there is no observed variation of initial phase as a function of period. Since the phase velocity is not as precisely known for periods greater than 300 sec as the shorter periods, the disagreement may be due to an error in phase velocity. Observations of initial phase for the southeast Alaska earthquake of July 10, 1958, for shorter periods also showed no definite variation of initial phase as a function of period [Brune, 1961]; however,

TABLE 6. Love Wave Phase Velocities

	<i>T</i>	<i>C</i>
A. Chile to Ñaña, Complete Great-Circle Path		
G_4-G_2	705	6.332
	636	6.159
	612	6.153
	552	5.982
	498	5.866
	282	5.203
G_6-G_4	690	6.205
	642	6.136
	582	6.030
	540	5.987
	528	5.950
	492	5.805
	492	5.810
	435	5.719
	360	5.423
	339	5.347
	330	5.330
	315	5.251
G_8-G_6	690	6.407
	615	6.179
	585	7.541
B. Chile to Isabella, Complete Great-Circle Path		
G_4-G_2	301	5.222
	277	5.138
	262	5.074
	198	4.874
	174	4.856
G_6-G_4	378	5.487
	366	5.432
	363	5.446
	342	5.338
	336	5.302
	297	5.109
G_8-G_6	396	5.591
	393	5.586
	390	5.579
	372	5.476
	359	5.392
	345	5.340
	330	5.279
C. Chile to Isabella, Complete Great-Circle Path		
G_6-G_2	315	5.236
	330	5.302
	372	5.469
	361	5.419
	333	5.291
	324	5.248
	285	5.083
D. Chile to Ogdensburg, Complete Great-Circle Path		
G_4-G_2	700	4.759

the radiation of energy was much more metric from the southeast Alaska earthquake than from the Chilean earthquake.

The mean initial phase in the period 200 to 550 sec is $0.02 \times 2\pi$ at Isabella, $0.59 \times 2\pi$ at Ogdensburg, that is, opposite within the limits of accuracy of the measurement, and hence Isabella and Ogdensburg probably lie in adjacent zones of the radiation pattern. The initial phase for R_2 at Ñaña for periods near 225 sec is the same as that at Ogdensburg in the same period range. The difference in azimuth to the two stations is only 5° . This indicates that a nodal line for the radiation pattern passes between the azimuths of Ñaña and Isabella, somewhat more west than the strike of the fault as indicated by the aftershocks. The single measurement of initial phase on the train R_3 at Isabella indicates the same initial phase as the train R_2 . Brune [1938] found that, for most of the observations, the initial phase on the Southeast Alaska earthquake, opposing azimuths at the source show the same phase. The model of the fault as a moving source, given above, would predict a phase difference of about 0.7π . The large variation of initial phase as a function of period at Isabella as well as the lack of difference between the initial phase of R_2 and R_3 suggest a point source as the mechanism. However, theories based on a point source do not predict the asymmetry of amplitudes observed for this earthquake.

Analysis of Love wave data. Table 5 gives data for Love waves from the seismograms. The Ñaña seismograms show a clear train of Love waves with periods extending from 700 sec to the G wave (period about 300 sec). The order of arrivals of this train G_4 , G_6 , and G_8 are also clearly shown on the record for the strain seismograph at azimuth $39^\circ 02'$. The G wave is not sufficiently dispersed to be analyzed for group velocity in the ordinary way and in fact, only been measured by *Satō's* method, which is dependent upon computing the group velocity from the phase velocity and its derivative. The group velocities for the waves preceding the G waves are shown in Figure 7. They agree well with theoretical calculations derived from free periods of oscillation [Pekeris, Alterman, and Jarosch, 1961; Landisman, and Ewing, 1960].

TABLE 7. Love Wave Phase Velocities between Rio de Janeiro (R_1) and Mt. Tsukuba (R_2)
Determined from the New Guinea Earthquake of 1958

T , sec	C , km/sec
150	4.790
120	4.708
110	4.679
100	4.648
90	4.617
80	4.577
70	4.528
60	4.474

more careful analysis to determine group velocity from the trains G_2 , G_4 , G_6 , and G_8 was carried out in a manner such that any influence of initial phase or instrumental phase shift was eliminated. The results are plotted as stars in Figure 7. The data from Table 5 are analyzed to determine the values of phase velocity for Love waves in Table 6. Tables showing the complete computations are filed with the American Documentation Service. The correlations of the Love wave velocity on the Ñaña record are clear and unambiguous. They give Love wave phase velocity data for the period range 700 to 300 sec. The change in calculated phase velocity for $T = 370$ sec

which would result from an error of one unit in assigning n is indicated in Figure 7. The change in phase velocity is quite large; consequently it is possible to demonstrate that proper choice of n has been made. Once the phase velocity is determined through a proper choice of n for one period, it is possible to extend it without difficulty to all other periods for which data are available. Since the precision in measurement of period is not critical for phase velocity determination by this method, observations of G waves may yield precise values of phase velocity for a limited period range.

From the Isabella record, the trains G_2 , G_4 , G_6 , and G_8 may be correlated to determine phase velocity. Periods are assigned to the various peaks by measuring the slope of a peak vs. arrival-time curve in the ordinary manner. Owing to the pulselike character of the G wave, there is serious uncertainty in the assignment of period, and the data cannot be used to define the slope of a group velocity curve; however, the data are best fitted if a small amount of normal dispersion is assumed. The phase velocities derived from the Isabella seismogram are shown in Table 6. They agree well with those found from the Ñaña record.

With the uncertainty in crest identification eliminated, we may use the published data of

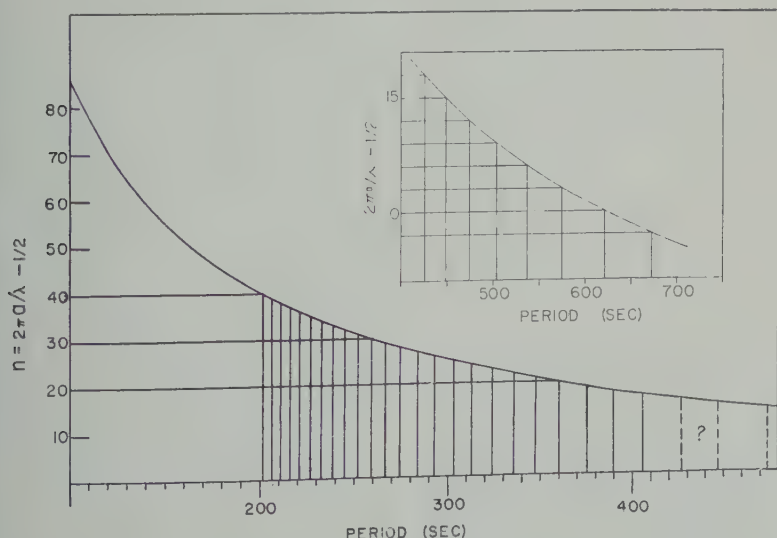


Fig. 8. $n = 2\pi a/\lambda - 1/2$, determined from phase velocity data in Figure 7 (excluding points from free oscillations). Integral values of n indicate the periods of free torsional oscillation of the earth.

TABLE 8. Torsional Free Periods Determined from Phase Velocities Compared with Observed Spectral Peaks

T, sec			
Phase Vel.	Pal. E-W	Isabella	Mode
	712		
673	673	672.6	$1T_9$
	633		
621		619.8	$1T_{10}$
574	579	576.0	$1T_{11}$
537	537		$1T_{12}$
	515		
504	500		$1T_{13}$
474	475		$1T_{14}$
	461		
449			$1T_{15}$
	435		
426			$1T_{16}$
	416		
406			$1T_{17}$
390	389		$1T_{18}$
	382		
375.2	374		$1T_{19}$
360.0	362		$1T_{20}$
	355		
347.6			$1T_{21}$
335.7			$1T_{22}$
324.1	325		$1T_{23}$
312.8	315		$1T_{24}$
303.3	306		$1T_{25}$
	298		
292.8			$1T_{26}$
283.9			$1T_{27}$
274.3			$1T_{28}$
266.4			$1T_{29}$
259.1			$1T_{30}$
252.2			$1T_{31}$
245.1			$1T_{32}$
238.7			$1T_{33}$
232.7			$1T_{34}$
226.8			$1T_{35}$
220.9			$1T_{36}$
215.9			$1T_{37}$
210.9			$1T_{38}$
206.3			$1T_{39}$
201.6			$1T_{40}$

Satô [1958], corrected for the polar phase advance, to extend the phase velocity curves to shorter periods. Figure 7 shows corrected data of Satô which was derived from *G* waves circling the earth. The great circles over which the correlations were made are mostly oceanic. The data scatter very little and extend the phase velocity curve down to periods below 100 sec. From the New Guinea earthquake of 1958, additional data give phase velocities in the

period range 60 to 250 sec as shown in Figure 7 and Table 7. These data agree closely with results of Satô and extend the period range measured Love wave phase velocities from 100 to 100 sec, a range which is comparable to for which Rayleigh wave dispersion data are known precisely. Theoretical phase and group velocities for the Gutenberg-Bullen A and Jeffries-Bullen earth structures are also shown in Figure 7. These models all have a continuous crust, and hence the observed phase velocity values may be expected to deviate at short periods. For periods greater than 200 sec the observed phase velocities agree generally with the Gutenberg-Bullen A model, but there are small deviations from the calculated curve which appear to be significant and systematic. Some very long period oscillations (20 to 30 min) appear ahead of the Rayleigh wave phase maximum at 2030 GCT on the record from the Ñaña seismograph, azimuth 309°, with group velocities high enough to suggest the very long period fundamental Love wave curve shown in Figure 9.

Love wave phase velocities and free vibration periods. Phase velocities were given above in the progressive wave analysis of the records from the Isabella, Ñaña, and Ogdensburg seismographs, from Satô's data from the Khatka and New Guinea (1938) earthquake and from Love wave data from the New Guinea earthquake of 1958. The period range measured is 60 to 700 sec. For periods within this range the observed phase velocities can be used to compute periods of free oscillations. In Figure 8 the quantity $n = 2\pi a/\lambda - 1/2$ is plotted as a function of period, λ being obtained from the phase velocity curve. The periods corresponding to integral values of n are free periods. Data from analysis of progressive waves are seen to determine periods of free torsional oscillations from $n = 86$ to $n = 9$. For n less than 40 or 50, large regional variations may be expected, but these variations cannot be established by the limited available data.

The periods read from Figure 8 are given in Table 8 and compared with the values obtained by Alsop, Sutton, and Ewing from the Palisades E-W pendulum seismograph of the Chilean earthquake and values obtained by Benioff, Press, and Smith [1961] by analysis of records of the Chilean earthquake from the Isabella

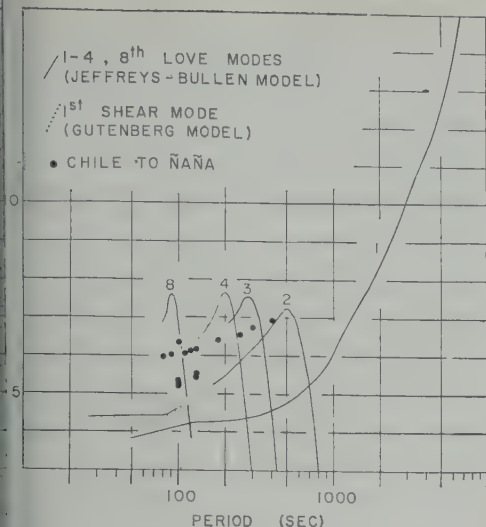


Fig. 9. Possible higher-mode group velocities obtained from the Ñaña seismogram for azimuth 02° .

in seismograph. The Palisades E-W seismograph should be sensitive only to Love waves from Chile. When additional free periods can be measured by spectral analysis in the period range, it will be of interest to compare them with those given in Table 8.

In Figure 7 the phase velocities derived from spectral peaks identified from the record of the Palisades E-W seismograph are indicated by solid triangles and found to agree very well with the phase velocity curve in the range of n from 20 to 26. Other spectral peaks given by Alsop, Eaton, and Ewing for lower-order n also agree with theoretical periods and periods derived from the observed phase velocities.

Possible higher modes. Waves of shorter period than the fundamental clearly appear ahead of G_2 on the Ñaña seismogram for azimuth 02° (see Table 9, filed with American Documentation Institute), and they also seem to appear on the Isabella and Ogdensburg seismograms. These data fit approximately the higher-mode Love wave curves given by Sató, Huisman, and Ewing [1960] on an extrapolation of the first shear mode curve of Bolt and Ewing (1961) as shown in Figure 9.

Conclusions.

Mantle Rayleigh waves have been identified in progressive trains of low order in the period range from 225 sec to about 1000 sec, and a group

velocity curve has been constructed for them, showing a range of group velocities from 3.55 to 7.8 km/sec, in agreement with the results reported by Brune, Ewing, and Kuo [1961].

2. Phase velocities for mantle Rayleigh waves have been determined by correlation of crests and troughs for successive path increments of 1 earth circumference. The experimental phase velocity curve is thus extended to 6.53 km/sec at a period of 650 sec.

3. Mantle Love waves have been identified in progressive trains of low order with group velocities of 4.4 km/sec at G to 5.0 km/sec at a period of about 700 sec on the Ñaña seismograms.

4. Phase velocities for mantle Love waves have been determined by correlation for successive path increments of 1 earth circumference. With Sató's data this extends the range of measured Love wave phase velocity from about 4.5 km/sec at 60 sec to about 6.45 km/sec at 720 sec.

5. As expected from the high values of group velocity attained for long-period Love and Rayleigh waves, both R_1 and G_1 occur as long-period displacements ahead of S at Ñaña ($\Delta = 26^\circ$) and Isabella ($\Delta = 83^\circ$).

6. Initial phases have been calculated, using the new phase velocity data, from the trains R_2 , R_3 , and R_4 at Isabella, and R_2 at Ogdensburg and Ñaña. Comparison of initial phases of Rayleigh waves at Ñaña and Isabella indicate that a nodal line passes between these two stations. Ogdensburg and Isabella also show opposite initial phases for Rayleigh waves in the period range 250 to 550 sec. Both results agree approximately with the strike of the fault indicated by aftershocks, but suggest a strike somewhat west of north.

At all three stations the even orders of surface wave trains are much stronger than the odd orders, a result for which Benioff, Press, and Smith have offered the explanation of a southward-traveling source.

7. The work presented here illustrates the power of analysis of seismograms in terms of progressive waves. In a strictly steady state condition, the analysis of a system in terms of progressive or of standing waves should give identical results. But in the presence of absorption, any advantage of the standing wave system disappears rapidly as the decrement of

the system increases. Analyzing a seismogram into a series of wave groups makes it possible to observe patterns which might be very difficult to detect in a spectral analysis.

8. The observed dispersion of Love and Rayleigh waves is in close agreement with theoretical dispersion based on the Gutenberg-Bullen A model of the earth. Small systematic deviations from the theoretical curve suggest that slight modifications of this model are necessary.

Acknowledgments. This work was supported in part by the Office of Ordnance Research under contract DA 30-069-Ord-2431, by the National Science Foundation, and by the Air Force Cambridge Research Center under project Vela.

REFERENCES

- Alsop, L. E., G. H. Sutton, and M. Ewing, Free oscillations of the earth observed on strain and pendulum seismographs, *J. Geophys. Research*, **66**, 631-641, 1961.
- Alterman, Z., H. Jarosch, and C. L. Pekeris, *Interim Report on Propagation of Rayleigh Waves in the Earth*, Dept. of Applied Math., The Weizman Institute, Rehovoth, Israel, 1961.
- Ben-Menahem, A., Radiation of seismic surface-waves from simple models of fault-planes, 1, Rayleigh waves, *Bull. Seism. Soc. Am.*, 1961.
- Benioff, H., and F. Press, Progress report on long period seismographs, *Geophys. J.*, **1**, 208-215, 1958.
- Benioff, H., F. Press, and S. Smith, Excitation of the free oscillations of the earth by earthquakes, *J. Geophys. Research*, **66**, 605-619, 1961.
- Bogert, B. P., An observation of free oscillations of the earth, *J. Geophys. Research*, **66**, 643-646, 1961.
- Bolt, B. A., and J. Dorman, Phase and group velocities of Rayleigh waves in a spherical, gravitating earth, *J. Geophys. Research*, **66**, 1961.
- Brune, J. N., Radiation pattern of Rayleigh waves from the southeast Alaska earthquake of July 10, 1958, *Symposium on Earthquake Mechanism*, *Publ. Dominion Observatory, Ottawa*, **24**, 10, 1961.
- Brune, J. N., M. Ewing and J. Kuo, Group phase velocities for Rayleigh waves of period greater than 380 seconds, *Science*, **133**(3455), 1958.
- Brune, J. N., J. E. Nafe, and L. E. Alsop, The polar phase shift of surface waves on a sphere, *Bull. Seism. Soc. Am.*, 1961.
- Dorman, J., M. Ewing, and J. Oliver, Study of shear velocity distribution in the upper mantle from mantle Rayleigh waves, *Bull. Seism. Soc. Am.*, **50**, 87-115, 1960.
- Ewing, M., and F. Press, Mantle Rayleigh waves from the Kamchatka earthquake of November 1952, *Bull. Seism. Soc. Am.*, **44**, 471-479, 1954.
- Ewing, M., and F. Press, An investigation of mantle Rayleigh waves, *Bull. Seism. Soc. Am.*, **45**, 127-147, 1954b.
- Gilbert, F., and G. J. F. MacDonald, Free oscillations of the earth, 1, Toroidal oscillations, *J. Geophys. Research*, **65**, 675-693, 1960.
- Jeans, J. H. The propagation of earthquake waves, *Proc. Roy. Soc. London A*, **102**, 554-574, 1923.
- Nafe, J. E., and J. N. Brune, Observations of phase velocity for Rayleigh waves in the period range 100 to 400 seconds, *Bull. Seism. Soc. Am.*, **50**, 427-439, 1960.
- Ness, N. F., J. C. Harrison, and L. B. Slichter, Observations of the free oscillations of the earth, *J. Geophys. Research*, **66**, 621-629, 1961.
- Pekeris, C. L., Z. Alterman, and H. Jarosch, Comparison of theoretical with observed values of periods of free oscillation of the earth, *Proc. Nat. Acad. Sci. U. S.*, **47**, 91-98, 1961.
- Satô, Y., Attenuation, dispersion and the waveguide of the *G* wave, *Bull. Seism. Soc. Am.*, **49**, 231-251, 1958.
- Satô, Y., Landisman, and M. Ewing, Love waves on a heterogeneous, spherical earth, 1, Theoretical periods for the fundamental and higher torsional modes; 2, Theoretical phase and group velocities, *J. Geophys. Research*, **65**, 2395-2404, 1960.
- Takeuchi, H., F. Press, and N. Kobayashi, Rayleigh wave evidence for the low-velocity zone in the mantle, *Bull. Seism. Soc. Am.*, **49**, 355-364, 1959.

(Manuscript received May 1, 1961;
revised June 7, 1961.)

Measurement of Q for Very Long Period Free Oscillations¹

LEONARD E. ALSOP, GEORGE H. SUTTON, AND MAURICE EWING

*Lamont Geological Observatory, Columbia University
Palisades, New York*

Abstract. The radical amplitude variation with time of long-period spectral peaks corresponding to the periods of free oscillation is explained. It arises from the rotation of the corresponding standing wave pattern with respect to the earth. The modulation caused by this rotation imposes certain conditions on the length of record and the interval between records used to measure Q by observing the amplitude decrement. Under these conditions the Q for the ${}_0S_2$ mode is measured to be 370, and for the ${}_0S_3$ mode, Q is measured to be about 300. These values are contrasted with other values of Q obtained by various authors for other free periods. Q appears to be constant for modes ${}_0S_2$ through ${}_0S_6$.

Introduction. It has been noted in a previous paper [Alsop, Sutton, and Ewing, 1961] connected with the spectral analysis of seismograms of the Chilean earthquake of May 22, 1960, that spectral peaks from two time intervals of 10 days varied radically in amplitude, failing to decrease with time according to any simple law. This puzzling phenomenon can be explained, at least for the modes of longest period, as occurring from the rotation, with respect to the earth, of the standing wave pattern of a given mode. The physical basis for the rotation of the standing wave pattern has been discussed by Backus and Gilbert [1961]. These authors have cautioned against determining the Q of free oscillations by using the apparent widths of spectral lines, as measured on the spectrum of a single observation, if the lines are not completely resolved. Q is defined in the usual manner by the amplitude decrement

$$A(t) = A(t_0) \exp(-\pi(t - t_0)/QT) \quad (1)$$

where $A(t)$ is the amplitude at time t of an observation with period T , and t_0 is the time of the first observation. However, determining Q directly from (1) by observing the amplitude decrement of a spectral peak obtained from observations at successively later times leads to meaningless results unless the record length used satisfies one of three conditions. Considering these conditions, we have determined the Q of

the lowest period mode of vibration of the earth, ${}_0S_2$, to be 370.²

Rotation of the standing wave pattern. C. L. Pekeris (personal communication, 1961) and Backus and Gilbert [1961] have calculated the splitting of the free oscillations of the earth resulting from the Coriolis force. Backus and Gilbert found that the frequency splitting is equal to $m\beta F$, where m is the degree of the tesseral harmonic describing the oscillations, F is the rotational frequency of the earth, and β is a parameter which for spheroidal oscillations depends both on the order of the Legendre polynomial and the number of nodes of the motion along a radius, and for torsional oscillations depends only on the order of the Legendre polynomial. These authors also show, using the addition theorem for Legendre polynomials, that the entire pattern of motion corresponding to a given order rotates relative to the earth's surface with a period equal to $1/\beta F$. From the values of β calculated by Backus and Gilbert for the ${}_0S_2$ and ${}_0S_3$ modes, it is found that these modes rotate with periods of 60.24 and 130.1 hours, respectively.

The necessity for the rotation of the pattern with this period may be shown by the follow-

²In our previous paper [Alsop, Sutton, and Ewing, 1961], we used the notation ${}_1S_2$ for this mode. Since then it has become the established practice to use a left subscript of zero rather than 1 to denote the fundamental mode, i.e., the left subscript refers to the number of nodal surfaces of constant radius for the given oscillation.

ing argument. Consider the motion $S_n^m(t)$, at a particular point on the earth's surface for a given mode described by P_n^m .

$$S_n^m(t) = \frac{A_m}{2} \cos [2\pi(f_0 + m\beta F)t] + \frac{A_m}{2} \cos [2\pi(f_0 - m\beta F)t] \quad (2)$$

where f_0 is the center frequency, and each frequency has equal amplitude $A_m/2$. These two frequencies will combine in the usual manner to give an amplitude and phase modulated wave.

$$S_n^m(t) = A_m \cos(2\pi m\beta F t) \cos(2\pi f_0 t) \quad (3)$$

with the frequency of modulation equal to $m\beta F$. The assumption that each frequency has equal amplitude is equivalent to stating that the motion is a pure standing wave. Unequal amplitudes will result in a standing wave plus a running wave in one direction. The modulation to be discussed is produced by the standing wave pattern only. The pattern of standing wave motion corresponding to degree m will have m nodal planes passing through the earth's axis. Therefore, if the pattern rotates with a period T , the amplitude at a given point will vary with a period T/m . From (3), T must equal $1/\beta F$, as before, a result which is independent of m . The expression (3) for a given order n will be of the form

$$S_n(t) = \sum_{m=0}^n A_m \cos(2\pi m\beta F t) \cdot \cos(2\pi f_0 t) \quad (4)$$

Criteria for measurement of Q . Consider two consecutive time intervals, each of length τ and with starting times differing by Δt , which are analyzed to obtain Fourier amplitudes of the free vibrations. Ignoring dissipative effects for the present, if $(\Delta t + \tau) \ll 1/n\beta F$ and $\tau > 1/f_0$, then from (4)

$$S_n(t) = \sum_{m=0}^n A_m' \cos(2\pi f_0 t) \quad (5)$$

where $A_m' = A \cos(2\pi m\beta F t_m)$, t_m being an appropriate average time in the time interval, and the analysis will yield essentially an amplitude for the frequency f_0 equal to $\sum_m A_m'$. Furthermore under these conditions the amplitudes determined for each record length will be

approximately equal. If $(\Delta t + \tau)$ is lengthy to approach $1/n\beta F$, the values of A_m' will differ in the two intervals unless $\Delta t \ll \tau$, and the amplitudes determined by analysis for these two intervals will differ. The amplitude of a given frequency in the second interval is as likely to be greater than that of the first interval as it is to be smaller.

Assume now that dissipative effects are of sufficient size to be observed in the length of record used. An attempt to measure Q over a given period by observing the decrement between two corresponding peaks in the two records cannot yield meaningful results for those modes having a value of $1/n\beta F$ such that $1/n\beta F$ is much less than $(\Delta t + \tau)$. With usual record lengths, i.e., of the order of several days, this effect will be of importance for only the longest period vibrations, because β decreases with increasing n . Also, probably for large values of n the patterns with large numbers of nodal planes though the axis will not be excited; i.e., A_m will be small for values of m of the order of n .

On the other hand, if $\tau > 1/\beta F > 1/f_0$, the lines will be completely resolved and the amplitude decrement between two records may be used to measure Q . However, Q may also be measured by measuring the line width [Berg and Smith, 1961].

One further possibility exists for the measurement of Q . If $\tau > 1/f_0$ and $\Delta t = k/\beta F$, where k is an integer, then each of the A_m 's will be averaged over the same region of the modulation cycle in both records, and the method of determining Q by observing the amplitude decrement may be used.

Thus, to measure Q , one of the three following conditions must be satisfied.

- A. $(\Delta t + \tau) \ll 1/n\beta F$, $\tau > 1/f_0$.
- B. $\tau > 1/\beta F$, Δt arbitrary.
- C. $\tau > 1/f_0$, $\Delta t = k/\beta F$ (k integer).

Condition A is usually satisfied for all but very long periods of free vibration, since the corresponding patterns rotate slowly. For very long periods either condition B or condition C must be satisfied to determine Q .

Determination of Q for S_2 and ${}_6T_2$. Successive digitized records from the Ogburn strain seismometer of length 20.35 sec

encing at 23h 49m 00s on May 22, 1960, analyzed for Fourier components by the method described by *Alsop, Sutton, and Ewing* [1961]. It will be noted that Δt between the third and fourth records is 61.05 hours, which is close to $1/\beta F$ for the ${}_0S_2$ mode. Therefore condition C of (6) is satisfied for these two modes. The amplitudes of ${}_0S_2$ and ${}_0S_3$ measured over four lengths of record are shown in Figures 1 and 2. Inserting the ratio of the amplitudes of the first and fourth peaks in (1) yields $Q = 370$ for ${}_0S_2$.

It has been assumed here that (1) may be satisfied by substituting the amplitudes obtained from the two spectra and setting $t = \Delta t = 61.05$ hr. The amplitude obtained in a Fourier analysis is an average of the amplitudes throughout the record length. If the records are of the same length, it is valid to substitute these average amplitudes directly into (1) [*Ness, Harris, and Slichter*, 1961].

The amplitudes observed for ${}_0T_2$ and ${}_0T_3$ are shown in Figures 3 and 4.³ The standing wave

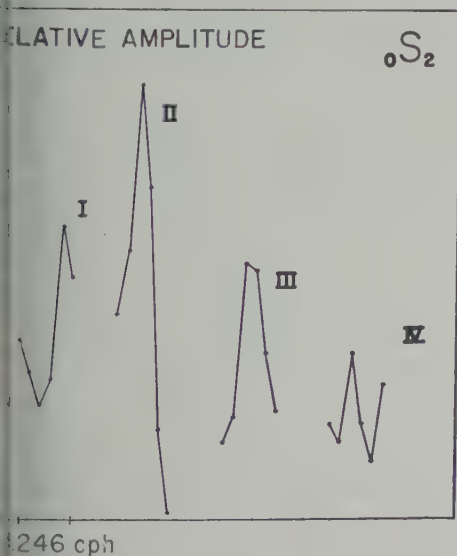


Fig. 1. Relative amplitudes of ${}_0S_2$ mode in four consecutive 20.35-hour intervals from strain seismogram (see text). Frequency interval between end points of peak indicated at bottom of figure.

³Donald and Ness [1961] have suggested that this peak may consist of both the ${}_0T_2$ and ${}_0T_3$ modes.

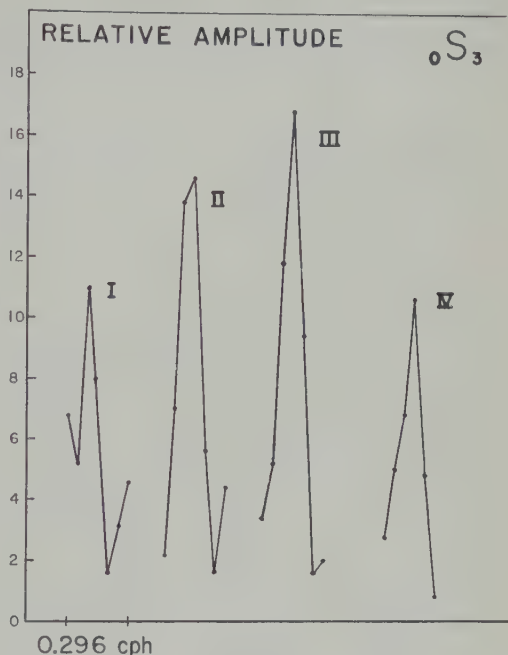


Fig. 2. Relative amplitudes of ${}_0S_3$ mode in four consecutive 20.35-hour intervals from strain seismogram (see text). Frequency interval between end points of peak indicated at bottom of figure.

patterns for these two modes rotate with periods of 12 and 30 days, respectively [*C. L. Pekeris*, personal communication, 1961; *Backus and Gilbert*, 1961]. Therefore $1/n\beta F$ equals 4 days for ${}_0T_2$ and 6 days for ${}_0T_3$. The record lengths used do not quite satisfy condition A of (6). However, this condition may be too stringent for ${}_0T_2$, at least, for it is probable that the amplitudes of the tesseral harmonics of highest degree may be small. Attempts to calculate Q for ${}_0T_2$ led to widely varying values, while for ${}_0T_3$ a Q of 300 satisfied the first three peaks.

Sources of error. It has been assumed in the determination of Q that the background noise level is negligible. The effect of the noise level on the ratio of two peaks is estimated to be less than ± 5 per cent, corresponding to a possible error of ± 30 for the Q of ${}_0S_2$. The Q 's given here have been calculated from the amplitudes illustrated in Figures 1 and 4. In our previous paper [*Alsop, Sutton, and Ewing*, 1961] the amplitudes and periods were obtained by parabolic interpolation. The precision of the Q 's given here is considered to be not sufficient, be-

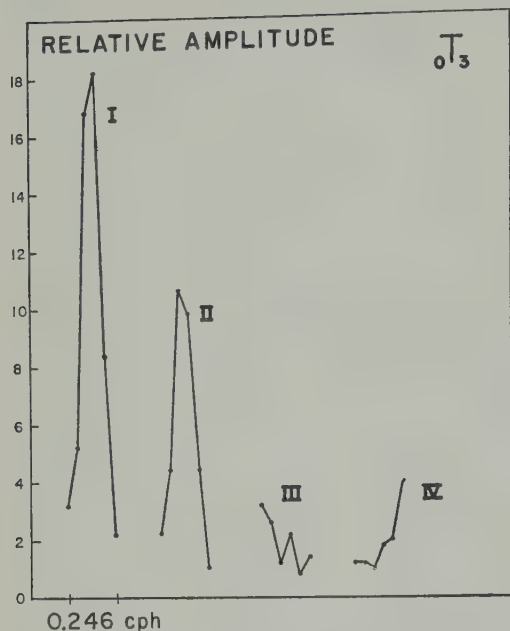


Fig. 3. Relative amplitudes of $0T_3$ mode in four consecutive 20.35-hour intervals from strain seismogram (see text). Frequency interval between end points of peak indicated at bottom of figure.

cause of the uncertain noise level, to justify this additional step.

In addition, for $0T_3$, as already mentioned, the value of Q obtained may be subject to some error, since it does not satisfy condition A of (6) exactly.

Conclusions. The striking phenomenon, illustrated in Figures 1 and 2, in which the amplitudes of the very long period free oscillations appear to increase with time can be explained as being due to the rotation of the standing wave pattern of the free oscillations with respect to the earth. Moreover, a simple study of the modulation of the amplitudes caused by the rotation of the pattern makes it possible to establish criteria for the record length and the interval between records, one of which criteria must be satisfied in order to obtain information about dissipative effects.

The value of Q for the $0S_2$ ($T = 53$ min) spheroidal mode is 370. Benioff, Press, and Smith [1961], who determined Q by measuring the width of the line, have found that the $0S_2$ ($T = 35$ min) mode has a Q of about 380. They were able to set a lower limit of 300 for the Q of $0S_2$

from the width of their spectral window.⁴ used approximately 11 days' record, or two periods of rotation for the $0S_2$ standing pattern. Therefore, the component frequency of $0S_2$ could be completely resolved, and Q measured by observing line widths. (They shown by these authors probably correspond to $m = -2$ and $m = +2$, with a suggestion $m = 0$. The peaks for $m = \pm 1$ and $m = 0$ appear to be missing.) Ness, Harrison, and Slichter [1961] have determined average Q for $0S_2$ ($T = 630$ sec) and $0S_{12}$ ($T = 500$ sec) as 366 and 280, respectively, by observing the amplitude decrement in four sections of record. It is not possible to apply the criteria of (6) to these measurements because the values of Q for $0S_2$ and $0S_{12}$ are not available. However, an indication of the reliability of these numbers is given by the relatively small variation of Q they determined for the different intervals used. It therefore

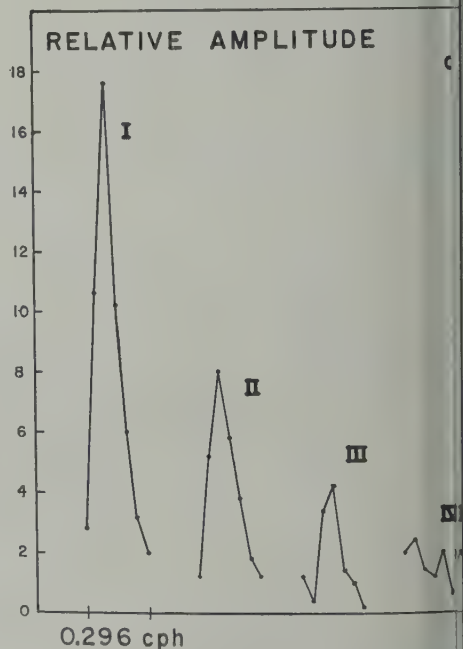


Fig. 4. Relative amplitudes of $0T_3$ mode in four consecutive 20.35-hour intervals from strain seismogram (see text). Frequency interval between end points of peak indicated at bottom of figure.

⁴ S. H. Smith, private communication, says they have now determined Q for the $0S_2$ mode to be 350, from a measurement of the width of the $m = 0$ line.

is likely that Q for the fundamental tidal modes of orders 2 through 9 is essentially constant, and is higher than Q for periods near 200 seconds as determined for Rayleigh waves by Ewing and Press [1954] and corrected by Brune [1961].

The value of Q for T_2 is measured to be about 10 times this value of Q is higher by a factor of 3 than the value found by Satô [1958] at 360 seconds for Love waves and corrected by Brune [1961]. Therefore, Q seems to be high for the higher-order fundamental torsional oscillations, just as for the low-order fundamental spheroidal oscillations.

As discussed previously, it has been assumed in the treatment that the amplitude is the same for modes of degree m and $-m$ for the same period. If this is not the case and if a running wave is also present, the modulated amplitude would be added to a constant amplitude due only to the running wave, thereby reducing the amplitude variations of the amplitude. In fact, the presence of the anomalous amplitude variation at a given period is indicative of the asymmetry of the radiation pattern of the source at that period. This is perhaps another reason why the value of Q could not be calculated for the T_2 mode although it did not satisfy condition A of section 2.

Acknowledgments. We wish to thank Professor Nafe and Y. Satô and Drs. J. Dorman and J. H. Van der Bruggen for several helpful and interesting discussions about the rotation of the patterns and the energy distributions of fundamental modes.

The Fourier analyses were performed numerically on an IBM 7090 through the courtesy of the General Electric Company, Syracuse, New York. The strain seismograph which recorded the data used is installed in a mine at Ogdensburg, N. J., belonging to the New Jersey Zinc Company.

This research was supported in part by the Air Force Cambridge Research Center.

REFERENCES

- Alsop, L. E., G. H. Sutton, and M. Ewing, Free oscillations of the earth observed on strain and pendulum seismographs, *J. Geophys. Research*, **66**, 631-641, 1961.
- Backus, G., and F. Gilbert, The rotational splitting of the free oscillations of the earth, *Proc. National Acad. Sci.*, to be published, 1961.
- Benioff, H., F. Press, and S. Smith, Excitation of the free oscillations of the earth by earthquakes, *J. Geophys. Research*, **66**, 605-619, 1961.
- Brune, J. N., Attenuation of dispersed wave trains, *Bull. Seism. Soc. Am.*, in press, 1961.
- Ewing, M., and F. Press, Mantle Rayleigh waves from the Kamchatka earthquake of November 4, 1952, *Bull. Seism. Soc. Am.*, **44**, 471-479, 1954.
- MacDonald, G. J. F., and N. F. Ness, A study of the free oscillations of the earth, *J. Geophys. Research*, **66**, 1865-1911, 1961.
- Ness, N. F., J. C. Harrison, and L. B. Slichter, Observations of the free oscillations of the earth, *J. Geophys. Research*, **66**, 621-629, 1961.
- Satô, Y., Attenuation, dispersion and the wave guide of the G wave, *Bull. Seism. Soc. Am.*, **48**, 231-251, 1958.

(Manuscript received April 12, 1961; revised June 14, 1961.)



Seismic Profiler¹

J. I. EWING AND G. B. TIREY

*Lamont Geological Observatory, Columbia University
Palisades, New York*

Abstract. Equipment for recording and profiling seismic refraction, seismic reflection, and underwater sound transmission data has been developed at Lamont Geological Observatory. The detector is a piezoelectric crystal hydrophone, the amplifiers are transistorized, modeled for those normally used in seismic refraction work, and the recorder is a modified Times Facsimile drum recorder. Profiling is achieved by initiating the drum rotation with the shot-instant signal. Refraction data are displayed in the form of a standard time-distance plot, the distance scale being determined by the speed of the shooting ship and the time scale by the speed of rotation of the drum, which can be preset by proper choice of gears and electrical clutching. Reflection data are recorded in section form, analogous to standard echo sounder records. A choice of full-wave or half-wave rectification is available, and logarithmic or linear amplifier response can be selected as desired.

INTRODUCTION

The seismic profiler² described here was developed to reduce the time required for analysis of seismic refraction and reflection data and to display the data in profile form so that interpretation of arrivals is facilitated. For refraction shooting, an entire profile, regardless of depth, can be recorded and displayed in a manner similar to the record of a geophone spread. The equipment is ideally suited to two-ship operation where it is simple to fire a series of shots but difficult to establish a line of receivers, particularly in deep water. For reflection shooting, the equipment can be operated for sub-bottom depth recorders [Knott and Hersey, 1956; Beckmann, Roberts, and Knott, 1959] with which spark or gas explosion sources are used, or it can be operated

with explosive charges when a more powerful source is required.

DESCRIPTION OF THE EQUIPMENT

Mechanical system. The recorder, a Times Facsimile drum recorder, has been used previously for precision echo sounding [Luskin, Heczen, Ewing, and Landisman, 1954]. The principal modification to this unit has been the installation of a friction clutch between the drive motor and the drum. This allows the drum to be stopped without stopping the motor. Between shots the drum is held at zero reference position, and the clutch slips. The shot-instant signal is amplified and releases the drum by means of a relay. Thus recording is commenced at the shot instant, and since the stylus traverses steadily with time, the arrivals are profiled.

Lamont Geological Observatory Contribution

In discussions with J. B. Hersey of the Woods Hole Oceanographic Institution, we have learned that the name 'seismic profiler' is being used by other research groups to designate equipment that operates either as a refraction or a reflection profiler with closely spaced pulses and automatic correlation of events. We consider the name to be suitable for general usage, because each system, whether it uses a spark, electromagnetic, or gas source, incorporates the same principles of recording and correlation.

Electrical system. The hydrophone, pre-amplifier, and amplifier used in this equipment have been described previously [Luskin and Koelsch, 1958]. The only modification required for the new application was to change the output of the amplifiers, which were designed to match a low-impedance galvanometer circuit, to a high impedance to match grid circuits in the recorder.

A block diagram of the equipment is shown in Figure 1. The hydrophone signal is recorded on the drum by two styli. In refraction work

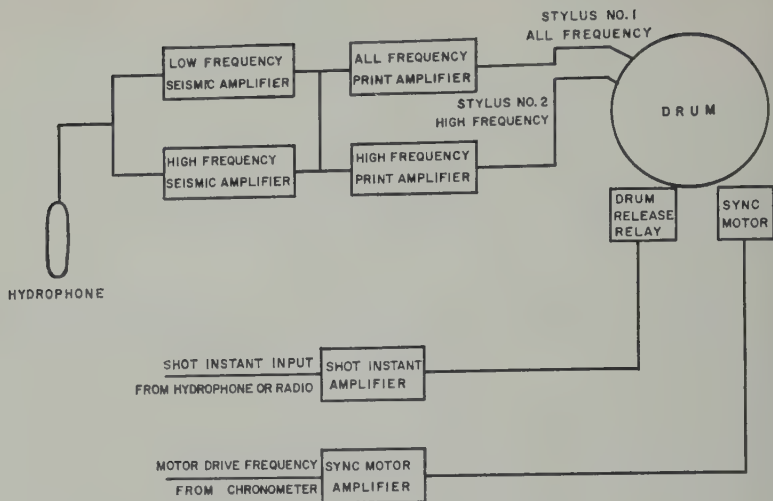


Fig. 1. Block diagram of seismic profiler.

one stylus records the output of a low-frequency channel of the seismic amplifiers, and the second stylus records only the signal from the high-frequency amplifier. Thus the first stylus records ground waves and water waves while the second records only water waves. It was found necessary to install the high-frequency stylus because strong sediment arrivals sometimes mask the water waves when both are recorded on a single trace.

The motor for the recording drum is driven by constant-frequency power furnished by a Westrex chronometer through a power amplifier. The drum is released to start the recording sequence by a shot-instant signal fed through a power amplifier to a relay.

The detailed wiring diagram of the recording unit and amplifiers is shown in Figures 2 and 3. J_1 is the input to the transistorized circuit for operating the drum release relay (R_1) at the shot instant. The relay can be operated also by a control switch (S_1) which has the following functions: (a) In the up (shot-instant) position, the shot-instant signal is fed to the relay circuit. (b) In the center (hold) position the shot-instant source is disconnected from the relay circuit to prevent keying by noise. (c) The down (test) position operates the relay to allow manual recording or to allow the drum to make more than a single revolution after having been initiated by a shot-instant signal.

J_2 connects the output of the shot-instant

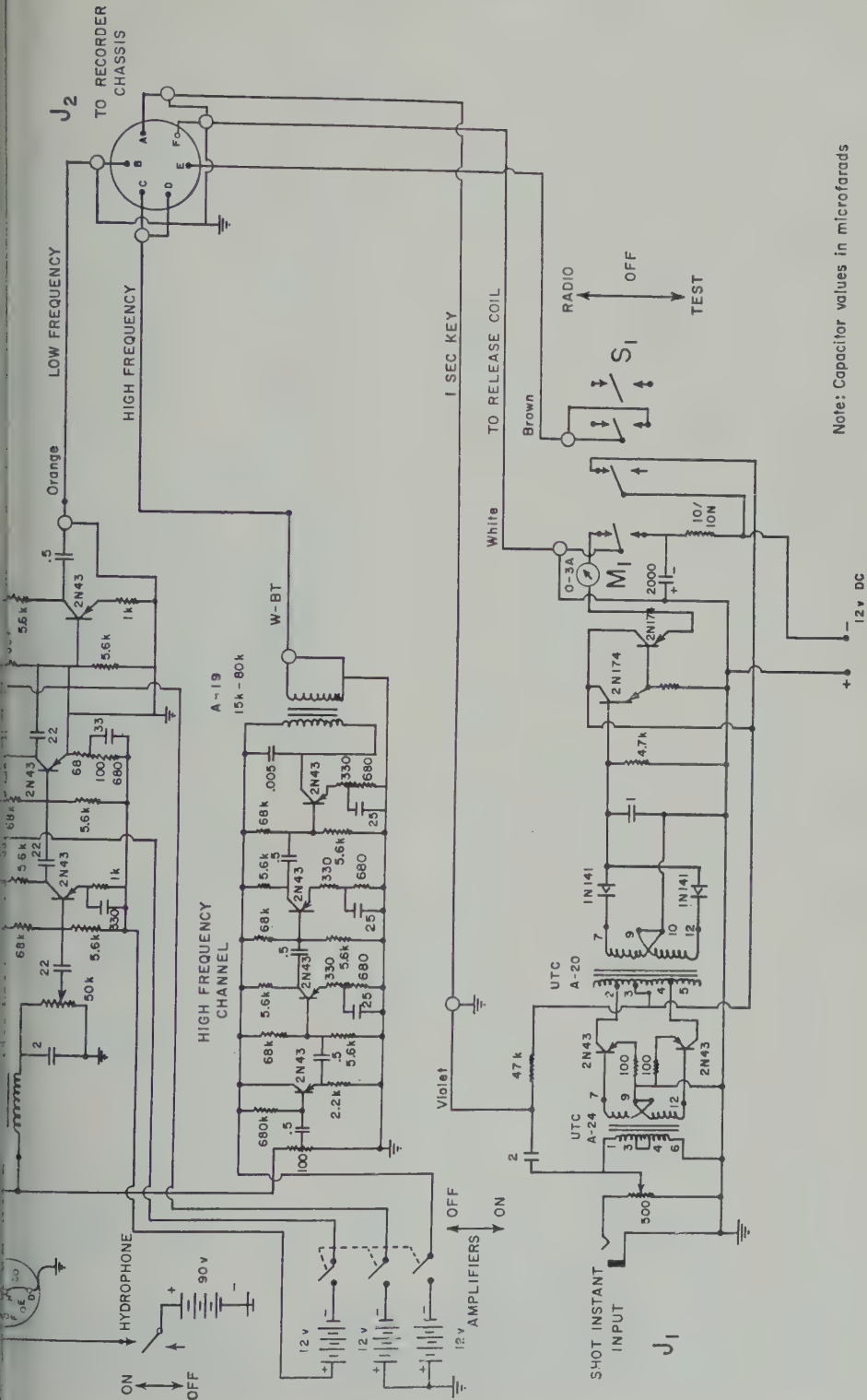
amplifier to the drum release relay and the output of refraction amplifiers to the input of the print circuit in the recorder. The low-frequency channel connection is made to switch S_2 which gives a choice of recording the low-frequency channel with either linear or logarithmic amplification. The high-frequency signals are always recorded linearly. Switch S_3 provides a choice of half-wave or full-wave rectification of the low-frequency signal. In the half-wave position the stylus prints only the positive signal peaks, and in full-wave position the stylus prints both positive and negative peaks. Refraction profile 2 (Fig. 4) was recorded with full-wave rectification; profile 3 with half-wave rectification (excepting shots 5, 6, and 7). The reflection profiles shown were recorded with full-wave rectification.

Switch S_4 is a microswitch which operates the drum release relay through switch S_1 when the drum is stopped. This prevents welding of the styli to the drum when the drum is stopped by signals received when the drum is in its normally stopped position.

Switch S_5 provides a choice of recording the high-frequency water wave with stylus 1 only or with stylus 2 only.

Switch S_6 allows timing marks to be recorded as small dots at the rate of 10 per second. The usual procedure is to record the timing marks between shots so that they do not interfere with the seismic arrivals.

The precision frequencies used for the



Note: Capacitor values in microfarads

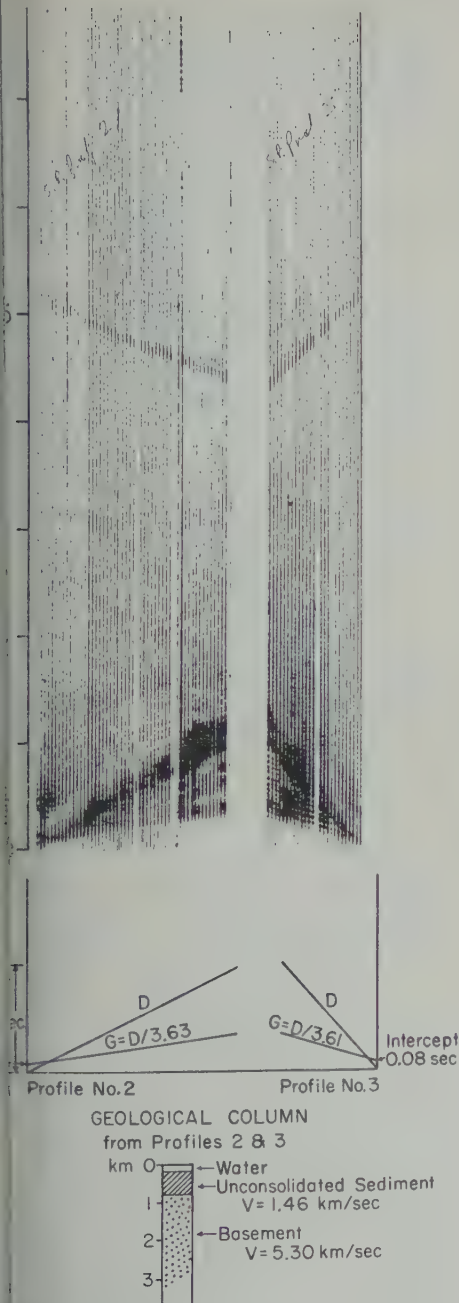
Fig. 2. Circuit diagram—amplifier section.

OPERATION AND MEASUREMENTS

Seismic refraction. For receiving seismic refraction data, it is essential to provide quiet listening conditions, as the arrivals are often low-amplitude, low-frequency (5 to 40 cps) signals. The techniques used to achieve the quiet listening conditions in marine operations are described in detail by *Officer, Ewing, Hen-ion, Harkrider, and Miller* [1959] and only a brief description will be given here. A hydrophone is suspended from the ship on a cable, along which floats are spaced to provide slight negative buoyancy. The cable is slacked before each shot so that the hydrophone sinks at a slow rate, is mechanically decoupled from the ship, and is at a sufficient depth to be below objectionable surface wave action. With the ship's engines and other machinery stopped the background noise is sufficiently low for good reception of the signals. For water depths not exceeding a few tens of fathoms, good results have been achieved by anchoring the receiving ship and laying the hydrophone on bottom.

Figure 4 shows two profiles recorded in preliminary field testing. These profiles were made before a separate high-frequency stylus and timing grid generator had been incorporated into the system. The receiving installation was made on the 70-ft vessel *Joseph Goldberger*, and a 12-ft outboard skiff was used as the shot boat. Following the standard two-ship refraction technique, the shot boat was equipped with a towed hydrophone and radio for transmitting the shot instants. The shots (blasting caps and seismograph boosters) were fired with burning fuse. The receiving ship was anchored and the receiving hydrophone placed on the bottom.

These profiles were received in Long Island Sound approximately 3 miles northwest of Lloyds Neck. Beneath the profiler records are shown the travel-time graphs traced from them and the computed structure section. The lines were not drawn on the records because of the desirability of displaying the data without the influence of interpretation. There is some deviation from exact alignment of the arrivals, largely caused in this case by variation in speed of the shot boat. Normally a larger vessel would be used for shooting, and the changes in slope caused by speed changes would not be noticeable at the scale of the records.



4. Record, travel-time graph, and velocity for refraction profiles 2 and 3 near Lloyds Long Island Sound.

us motor are brought from the Westrex meter through plug J . Switch S_7 provides of 5 or 10 seconds per revolution of m by changing the power frequency to ve motor.

In these profiles it is difficult to pick the exact time of arrival of the water waves. This is the result of mixing the high-frequency with the low-frequency amplifiers on a single stylus. When a strong low-frequency ground wave arrives coincidentally with the water wave it is impossible to identify either. It was for this reason that the modification was made to provide an additional stylus which records only high-frequency (water wave) arrivals. In profiles 2 and 3 (Fig. 4), the lines indicated as direct water waves (*D*) may actually be sediment lines, in which case the sediment velocity and basement velocity shown are low by a few per cent. Profile V-17-5 was recorded after the high-frequency stylus was installed. The high-frequency trace is displaced slightly to the right and considerably behind (approximately 1 second of drum rotation) the low-frequency trace. Figure 5 shows the profiler record and the travel-time graph and structure section derived from it. On the travel-time graph are plotted arrivals read from standard oscillograms which were made at the same time. A further comparison of the two methods of recording is shown in Figure 6 in which a low-frequency trace from an oscillogram is displayed below a profiler trace of the same shot. Because the time scales are slightly different, lines have been drawn to show the correlation of the peaks and troughs with the spots on the profiler trace.

The structure derived from the profiler records in Long Island Sound is in good agreement with that given by *Oliver and Drake* [1951] from the results of previous work in this area using a standard recording technique and with measurements made with a spark-source subbottom depth recorder [*Beckmann, Roberts, and Lusk*, 1959].

The strong late arrivals (4 to 5 seconds) in profiles 2 and 3 are reflected waves from a region south of the receiving position, probably near Oak Neck. They are low-frequency arrivals (< 30 cps), as can be determined from two shots in profile 2 which were recorded with high-frequency response only and failed to show the reflector. Because of the low frequency of these arrivals, it would be reasonable to suppose that the reflector is a subsurface feature: either an irregularity in the basement surface or possibly the contact of the coastal plains sediments with the basement. From the

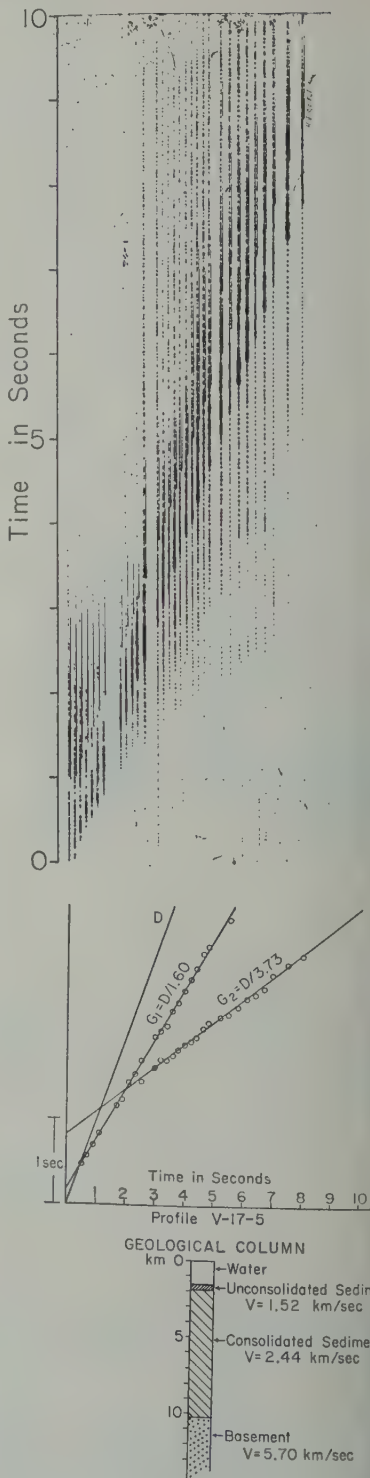


Fig. 5. Record, travel-time graph, and velocity section for refraction profile V-17-5 on the Florida shelf.

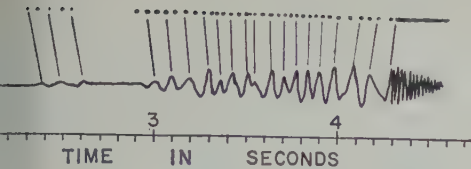


Fig. 6. Seismic profiler trace (upper) compared with oscillogram trace (lower). Profiler trace was with logarithmic amplification.

ational data, however, it appears that the line is the reflector, and the fact that high frequencies are not observed in the reflected signals could be attributed to greater attenuation along the transmission path and less velocity at the shore.

Other less obvious reflections can be seen at approximately the same alignment, indicating other reflectors south of the receiving station but less distant than the prominent reflection.

Seismic reflection. When used for reflection measurements, the seismic profiler produces a record similar in appearance to a standard echo sounder record or to those made by the various bottom depth recorders which have been used in recent years in shallow water areas [Lusk and Hersey, 1956; Beckmann, Roberts, and Lusk, 1959; Ewing, Lusk, Roberts, and Lusk, 1959; Moore, 1960]. Sample records made by the seismic profiler in deep water in the Yucatan Basin of the Caribbean Sea and in the Gulf of Mexico are shown in Figures 7 and 8.

For seismic refraction or reflection measurements a prime requisite is to provide a quiet hydrophone. This has been best achieved in the past by slacking the hydrophone cable between each shot. In reflection shooting the ship usually proceeds at 6 to 8 knots. Owing to the necessity of shooting at relatively close time intervals, the slacking operation has been mechanized. Figure 9 shows the arrangement used with success aboard *Vema*. In the upper part of the figure the hydrophone is pulled close to the ship before slacking, a bight of cable being held astern. The lower sketch shows the hydrophone cable slacked.

Charges larger than $\frac{1}{2}$ lb of TNT were used for the work reported here, although larger charges have been used when small ones were available. These were fired on a 2-minute

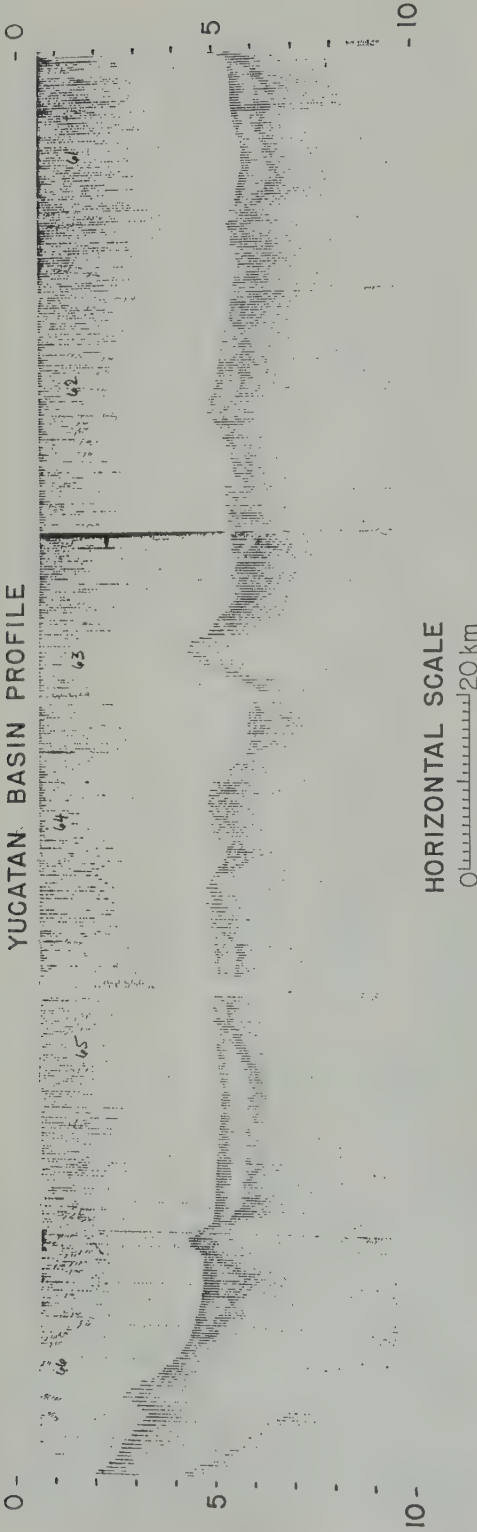
schedule in recording the profiles. The vertical exaggeration in these profiles is approximately 8 to 1. Obviously this is not a constant exaggeration, as the vertical scale changes with increasing depth owing to the velocity increase.

Figure 7 shows a profile which runs from the ridge near Grand Cayman Island to a point in the northwest edge of the Yucatan Basin. This profile was made with water works boosters (equivalent to $\frac{1}{18}$ lb of TNT). Maximum penetration of the bottom was about 1 km, corresponding to a delay time of approximately 1 second. The only refraction data in this area [Ewing, Antoine, and Ewing, 1960] is a profile in the basin south of the Isle of Pines which shows about 1 km of low-velocity sediment (2.1 km/sec) overlying a layer with a velocity of 3.9 km/sec. The depth of the interface between these two layers is closely comparable to the average depth of the deepest reflector in the reflection profile. The refraction data did not show the intermediate interfaces within the sedimentary column which are clearly seen in the reflection records.

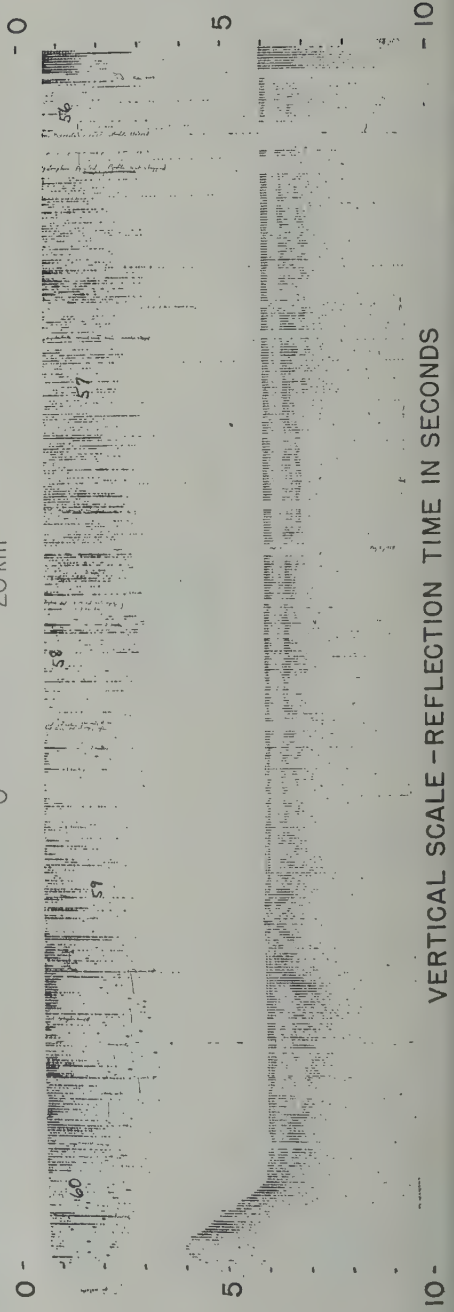
Figure 8 shows a reflection profile in the Gulf of Mexico running from the Campeche escarpment northwestward to a knoll southeast of the Sigsbee knolls [Ewing, Ericson, and Heezen, 1958]. Maximum penetration of the bottom here was about 5 km (approximately 4 seconds delay time) just south of the knoll. Refraction data in this general area [Ewing, Antoine, and Ewing, 1960] show layers with velocities 2.1, 3.2, and 3.8 km/sec overlying one with velocity about 5 km/sec. The deepest horizon, the top of the 5 km/sec layer, is seen more clearly than the shallower ones in the major portion of this profile, principally because it was necessary to use high recording gain to resolve the deep arrivals. This had the effect of over-printing much of the record in which the shallower reflections appear. As will be discussed later, provision is being made to print two records simultaneously at different gain levels so that shallow reflections will be preserved while sufficient gain is provided for recording deep ones.

The shots recorded in the profiles shown in Figures 7 and 8 were detonated at a depth of 30 to 40 feet, hence all show 'bubble pulse' caused by oscillation of the gas bubble. The interval between the first and second pulse is approximately .08 sec in the Yucatan Basin pro-

YUCATAN BASIN PROFILE



HORIZONTAL SCALE
0 20 km



VERTICAL SCALE - REFLECTION TIME IN SECONDS

Profile in the Yucatan basin of the Caribbean Sea from Cayman ridge near Grand Cayman Island (upper left) to northwestern

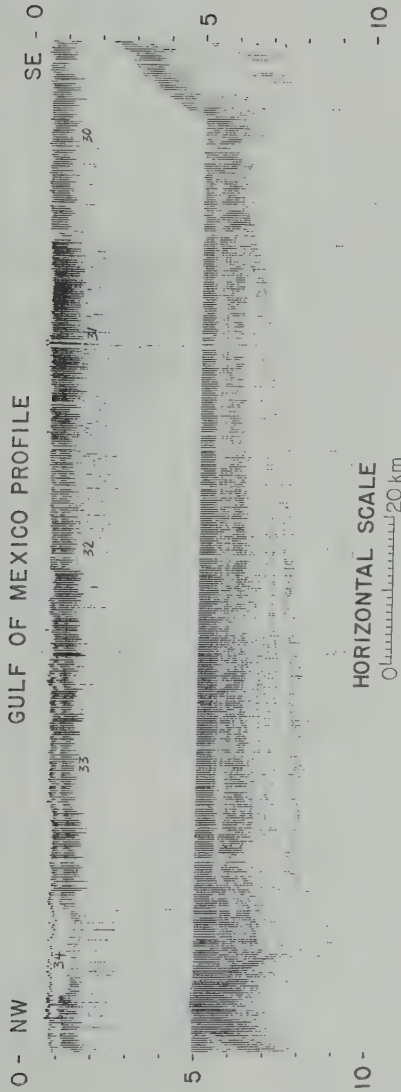


Fig. 8. Reflection profile in Gulf of Mexico from Campeche escarpment (right) to a knoll southeast of Sigsbee knolls (left). Water depth in basin is about 2000 fm. Deepest penetration of the bottom is 12,000 to 15,000 ft near the knoll. Deepest reflector is top of 5-km/sec layer [from refraction data given by *Ewing, Antoine, and Ewing, 1960*].

file and approximately .12 sec in the Gulf of Mexico profile, corresponding to water works boosters used in the former area and $\frac{1}{2}$ -lb charges of TNT used in the latter.

The results achieved here—continuous mapping of strata to a total reflection time of 9 seconds, using a single detector and recording channel from a ship underway—were accomplished through the application of several fundamental principles:

(a) Small explosions were used, at a depth chosen to reduce the prolongation of signal due to bubble pulses and surfaces echoes, as far as possible; (b) the background noise was reduced by placing the hydrophone a considerable distance from the ship and spacing buoys on the cable to permit the proper slow sinking of the hydrophone when the cable was slacked; (c) signal amplitude, the simplest index of reflection, was used as an indication of a reflector; (d) closely spaced shots were used to provide good data density; (e) the data were displayed in cross section form to facilitate correlation.

Modifications of the equipment. Some possible modifications that may be desirable for certain operations are listed below.

1. When using the recorder for reflection profiling it would be advantageous to have two levels of recording in order that the shallow sub-bottom reflections are not obliterated when the gain is set high enough to record deep reflections. One proposed plan is to have two styli, one low gain and the other high gain, separated by $\frac{1}{2}$ the width of the recording paper. Thus one half of each sheet of record will be printed at high gain and the other half printed simultaneously at low gain. For construction of long profiles, the sheets can be split, high gain halves being joined to form a high-

gain profile and low-gain sheets joined to form a low-gain profile. In addition to printing gain levels, all shots will be recorded on magnetic tape with at least 60-db dynamic range to provide the possibility for playback and analysis with different filter and gain settings.

2. When operating the profiler for reflection shooting, the styli can be translated at a rate proportional to the ship's speed by a motor controlled by the pit log. This will make the horizontal scale proportional to distance rather than to time.

3. A chart recorder can be used in place of a drum recorder to give a continuous profile on one sheet of paper.

4. For two-ship refraction operations, each ship can be equipped with a seismic profiler, each unit having a set of firing contacts and a drum for detonating charges electrically. The two units would run continuously during a profile, hence, the two drums will be in constant phase relationship to each other, and the closing of firing contacts on the shooting ship (zero time) will correspond to a specific time (drum rotation) at the recording unit. This method of synchronizing shot instants and recording has the advantage of not requiring the radio. The very accurate synchronization can be maintained with these recorders which are driven by tuning-fork-controlled oscillators, accurate to a few parts per million. From past experience it is estimated that the synchronization of units should not vary by more than .01 seconds per hour.

SUMMARY

The equipment and techniques described in this report enable us to make underway measurements of the thickness of the sedimentary cover in deep water areas of the oceans. Thus it is possible to produce an echogram which shows not only the topography of the ocean floor but also that of the underlying hard rock floor. In most places sampled to date, which include several hundreds of miles of track in the Gulf of Mexico, North Atlantic Ocean, Caribbean Sea, and southeastern Pacific Ocean, the measurements have reached the upper surface of the layer in which the seismic velocity is known to change from refraction data, to be in the range of 5.5 km/sec. The thickness of the sedimentary cover found in these areas was as much as

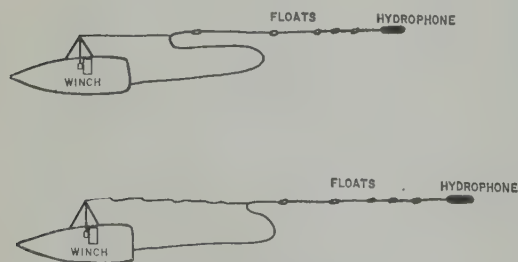


Fig. 9. Arrangement for slacking hydrophone cable with bathythermograph winch.

In addition to total thickness of the low-velocity sediments, much information has been gained about the stratification within the sedimentary column. Particularly in the Gulf of Mexico, where the sediments are thick, many features have been observed, showing horizon-parallel layering above and gently dipping beds below, somewhat deeper.

Seismic refraction measurements have shown a layer with velocity comparable to that of the deepest reflector in the profiles shown here is widespread, if not universal [Raitt, 1956; Bell, Hill, and Swallow, 1958; Ewing and Moore, 1959] in all the oceans. Whether it is the same layer everywhere is not known, nor is its composition known. Several writers have suggested that it may be consolidated sedimentary material; other have proposed volcanic or igneous rock. Whatever its nature, it is undoubtedly one of the most significant boundaries in the structural column, and the ability to map it by reflection measurements while underway will surely prove to be of considerable benefit in the over-all study of earth structure. Almost everywhere in the oceans sedimentary structure has been masked, at least partially, the relief of high-velocity crustal layers. In the abyssal plains, where large volumes of sediments have been brought in by the turbidity current, complete burial of large-scale topography has resulted. The ability to penetrate the sediments and measure the sub-bottom topography is an advance in the study of crustal structure comparable to the introduction of deep-sea sounders.

Acknowledgments. This work was made possible through the support of the Bureau of Ships, U.S. Navy, under contract NObsr 64547 and 85077, and the Office of Naval Research, contract Nonr 266(48), of the United States Navy.

It is a pleasure to acknowledge major contributions to this work by Albert Stockel, Donald E. Sheldon Gerber, and J. L. Worzel. Assistance in the field testing were P. Centeno, W. Pittman, H. Houtermans, C. Windisch, R. Marsh, and J. Zaiane of Columbia University, and J. Zaiane Yanez of the University of Mexico.

REFERENCES

- Beckmann, W. C., A. C. Roberts, and B. Luskin, Sub-bottom depth recorder, *Geophysics*, **24**, 749-760, 1959.
- Ewing, J., J. Antoine, and M. Ewing, Geophysical measurements in the western Caribbean Sea and in the Gulf of Mexico, *J. Geophys. Research*, **65**, 4087-4125, 1960.
- Ewing, J., and M. Ewing, Seismic refraction measurements in the Atlantic Ocean basins, in the Mediterranean Sea, on the mid-Atlantic ridge, and in the Norwegian Sea, *Bull. Geol. Soc. Am.*, **70**, 291-318, 1959.
- Ewing, J., B. Luskin, A. Roberts, and J. Hirshman, Sub-bottom reflection measurements on the continental shelf, Bermuda banks, West Indies arc, and in the west Atlantic basins, *J. Geophys. Research*, **65**, 2849-2859, 1960.
- Ewing, M., D. Ericson, and B. Heezen, Sediments and topography of the Gulf of Mexico, in *Habitat of Oil*, *Am. Assoc. Petrol. Geol.*, 995-1053, 1958.
- Gaskell, T., M. Hill, and J. Swallow, Seismic measurements made by H. M. S. *Challenger* in the Atlantic, Pacific and Indian Oceans, and in the Mediterranean Sea, 1950-1953, *Phil. Trans. Roy. Soc. London*, **251**, 23-83, 1958.
- Knott, S. T., and J. B. Hersey, High-resolution echo sounding techniques and their use in bathymetry, marine geophysics and biology, *Deep-Sea Research*, **4**, 36-44, 1956.
- Luskin, B., B. Heezen, M. Ewing, and M. Landisman, Precision measurements of ocean depth, *Deep-Sea Research*, **1**, 131-140, 1954.
- Luskin, B., D. Koelsch, Transistorized seismic equipment for refraction measurements at sea, *Lamont. Geol. Obs., Tech. Rept. 17*, CU-46-58, NObsr 64547, *Geol.*, 1958.
- Moore, D. G., Acoustic-reflection studies of the continental shelf and slope off southern California, *Bull. Geol. Soc. Am.*, **71**, 1121-1136, 1960.
- Officer, C., J. Ewing, J. Hennion, D. Harkrider, and D. Miller, Geophysical investigations in the eastern Caribbean: Summary of 1955 and 1956 cruises, in *Phys. and Chem. of the Earth*, **3**, Pergamon Press, 17-109, 1959.
- Oliver, J. E., and C. L. Drake, Geophysical investigations in the emerged and submerged Atlantic coastal plain, 4, Long Island area, *Bull. Geol. Soc. Am.*, **62**, 1287-1296, 1951.
- Raitt, R. W., Seismic refraction studies of the Pacific Ocean basin, 1, Crustal thickness of equatorial Pacific, *Bull. Geol. Soc. Am.*, **67**, 1623-1640, 1956.

(Manuscript received May 3, 1961.)

The Effect of Plasticity on Decoupling of Underground Explosions

A. L. LATTER, E. A. MARTINELLI
J. MATHEWS, AND W. G. MCMILLAN

The Rand Corporation, Santa Monica, California

Abstract. The effect of plasticity, including work hardening, on seismic decoupling of underground explosions has been studied for large spherical cavities designed to give maximum decoupling and for small (overdriven) cavities designed to give partial decoupling. An important result is that plasticity plays no role in explosions in large cavities, even those at great depth for which some plastic flow occurs during construction of the cavity. For small cavities at great depth plasticity affects the decoupling factor by an amount depending upon the degree of overdriving, the depth, and the detailed stress-strain relation of the medium. A further result is that for cavities at a depth of about 1 km in a medium like salt, which exhibits work hardening, the decoupling factor is at least as great as that obtained in the overdriven Cowboy experiments and could be appreciably greater. It is also concluded that the depth of the cavities in the Cowboy experiments was not sufficient for plastic flow to occur in the salt medium before other inelastic behavior, e.g., cracking, set in. A simplified model of the Cowboy inelasticity is discussed.

Introduction. A method for reducing the seismic signal from an underground nuclear explosion by setting off the explosion in a large underground cavity was proposed early in 1959 and later presented at the Geneva Nuclear Test Conference [Latter, 1959]. The theoretical basis for the method has recently been published [Latter, LeLevier, Martinelli, and McMillan, 1961] along with the results of the Cowboy experiments—high-explosive detonations in salt domes—which established the validity of the method [Murphey, 1961; Herbst, Werth, and Murphey, 1961].

The present paper grew out of a number of questions concerning the influence of plasticity on the seismic signal from decoupled explosions. It is well known that surrounding cavities at great depth, where the overburden pressure is comparable to the yield stress of the medium, are in a zone in which plastic flow occurs during the making of the cavity. The question arises: How does the presence of this plastic flow affect the decoupling factor?

Another question arises in connection with the cavity volume required for maximum decoupling. In salt, at a depth of about 1 km, the volume has been determined to be about 7×10^4 m³/kt from the requirement that the medium respond elastically to the explosive loading [Latter, 1959].

There are two kinds of inelasticity that are important: cracking and plastic flow. To avoid

making cracks, the hoop stresses must remain compressive because rock-like materials have little or no strength in tension. Latter [1959] has shown that this requires the average pressure in the cavity to be less than 3 times the overburden pressure. If the medium is already cracked, the average pressure in the cavity must be less than the overburden pressure itself in order to keep the cracks from opening up and propagating. Finally, to avoid plastic flow, stress differences must not become large compared with the yield stress of the medium, even during the passage of the large pressure pulse associated with the reflection of the shock wave from the cavity wall. This latter effect—plasticity due to the pressure spike—imposes the most stringent limit and fixes the volume quoted above. (At a depth of a kilometer the volume requirement of 7×10^4 m³/kt is equivalent to requiring that the average pressure in the cavity not exceed one-half of the overburden pressure.) However, this volume could clearly be reduced if the pressure spike were eliminated, as in fact it can be by methods already suggested [Latter, LeLevier, Martinelli, and McMillan, 1961]. In this case, there would be a new limit on the volume due to the plastic flow that is associated with the average pressure in the cavity. The question then arises: At what depth does this limit become more stringent than the cracking limit?

Finally, we note that the Cowboy experiments

were conducted at a depth of ~ 250 m, where the overburden pressure is only 50 bars. Therefore, even at 5 times the overburden—the greatest average cavity pressure in the Cowboy experiments—the pressure barely exceeds the elastic limit of the salt, and plasticity plays essentially no role in reducing the decoupling. However, for the nominal cavity at a depth of 1 km, 5 times the overburden pressure is ~ 1000 bars, which exceeds the elastic limit of salt [Handin, 1953], and therefore plastic effects cannot be ignored. The final question is: How should the results of the overdriven Cowboy shots be extrapolated to deeper cavities?

Stress and strain distribution before the explosion. It is convenient to state the equations in a form that applies to both plasticity and elasticity. Throughout we make the infinitesimal strain approximation.

The stress equilibrium equation, assuming spherical symmetry, is

$$\frac{\partial \sigma_r}{\partial r} + \frac{2}{r}(\sigma_r - \sigma_t) = 0 \quad (1)$$

where σ_r is the radial and σ_t the tangential stress. The stresses σ_r and σ_t are related to the displacement ξ (which is purely radial) in the following way:

$$\frac{1}{3}(\sigma_r + 2\sigma_t) + p_0 = k \left(\frac{\partial \xi}{\partial r} + \frac{2\xi}{r} \right) \quad (2)$$

where p_0 is the lithostatic pressure and k , the incompressibility, is assumed to be constant and the same for both elastic and plastic deformations. Equation 2 states that the incompressibility times the dilatation is equal to the change in average stress.

In addition, it is assumed that the shear stress $\sigma_r - \sigma_t$ is a function of the shear strain ϵ only, i.e.,

$$\sigma_r - \sigma_t = \sigma(\epsilon) \quad (3)$$

where the shear strain is the difference between the radial and tangential strains:

$$\epsilon \equiv \partial \xi / \partial r - \xi / r \quad (4)$$

In the case of elasticity $\sigma(\epsilon)$ is given by

$$\sigma(\epsilon) = 2\mu\epsilon \quad (5)$$

where μ is the modulus of rigidity. Since for small strains the medium is elastic, any actual function $\sigma(\epsilon)$ must approach $2\mu\epsilon$ near the origin.

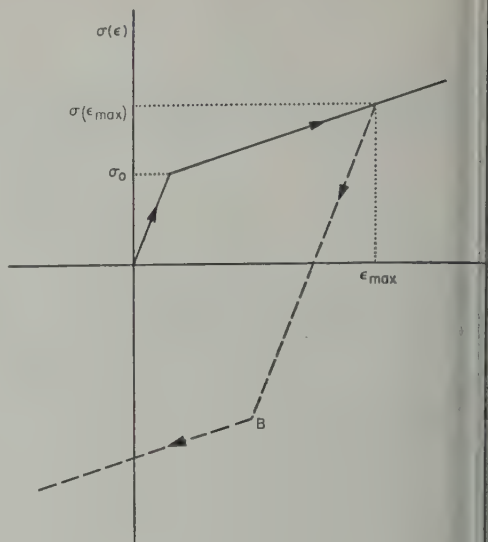


Fig. 1. The approximate stress-strain relation given by equations 6 and 7. The solid portion applies to unworked material. The dashed line to the point B shows a typical unloading path. Beyond B plasticity sets in again.

It should also be noted that (3) implies that the shear strain is independent of the average stress, which is only roughly true for real materials.

It is mathematically convenient to approximate the stress-strain relation $\sigma(\epsilon)$ by a piecewise straight-line segments as shown in the lines of Figure 1. In this approximation,

$$\sigma_r - \sigma_t = \begin{cases} 2\mu\epsilon, & (\sigma_r - \sigma_t \leq \sigma_0) \\ 2\alpha\epsilon + \sigma_0(1 - \alpha/\mu), & (\sigma_r - \sigma_t \geq \sigma_0) \end{cases}$$

where σ_0 is the elastic limit and α is a constant less than μ , chosen to give a reasonable fit to the measured stress-strain relation $\sigma(\epsilon)$. The approximation (6) exhibits the essential elastoplastic features of our problem and includes plasticity ($\alpha = 0$) as a special case.

Equation 3 applies for processes in which the shear stress does not decrease. The formation of a cavity by excavation or washing is such a process. However, for an explosion, which involves increasing the pressure in the cavity and therefore decreasing the existing shear stress, equation 3 is not applicable. In this case experience indicates that the shear stress decreases along a path parallel to the elastic portion of the stress-strain curve.

e, so that

$$\sigma_r - \sigma_t = 2\mu(\epsilon - \epsilon_{\max}) + \sigma(\epsilon_{\max}) \quad (7)$$

The μ is the same modulus of rigidity as before. ϵ_{\max} is the maximum strain reached by the medium at the location in question. Equation 7 is a simple description of the work-hardening process. It means that the medium behaves elastically under stress reduction but remains in a permanent strain state. The dashed line in Figure 1 represents (7), with arrows indicating the loading path. Note that plasticity begins again at the break point *B* where $\sigma = -\sigma(\epsilon_{\max})$.

To calculate the stresses resulting from the existence of a hole, we must solve the above equations subject to appropriate boundary conditions. We assume that the radius a of the hole is small compared with its depth and that the medium is in the medium before the existence of the hole is hydrostatic and equal to $-p_0$, that is, the negative of the lithostatic pressure at that point. It follows that far from the hole the stress is hydrostatic and equal to $-p_0$, i.e.,

$$\sigma_r = \sigma_t = -p_0, \quad (r \gg a) \quad (8)$$

and condition is that at the boundary of the (empty) hole the radial stress must be equal to zero, i.e.,

$$\sigma_r = 0, \quad (r = a) \quad (9)$$

We assume that initially the medium is everywhere at the origin of the stress-strain curve, i.e., there are no shear stresses or shear strains at the formation of the hole. In forming the hole each point of the medium moves along the stress-strain curve $\sigma(\epsilon)$ to some final state ϵ_{\max} which depends upon the radius r of the point and the hydrostatic pressure p_0 .

A maximum shear stress $(\sigma_r - \sigma_t)$ occurs, of course, at the cavity surface $r = a$. If the lithostatic pressure p_0 is less than some critical value, the shear stress at $r = a$ is less than σ_0 , and the entire medium is elastic. For p_0 greater than the critical value there is a radius $\rho(>a)$ such that the shear stress is exactly σ_0 , beyond which the medium is elastic and inside which the shear stress is determined by the plastic portion of the stress-strain curve. Requiring stresses and displacements to be continuous at the radius $r = \rho$ completes the boundary conditions.

The algebra is complex and uninteresting, but it leads to the following equation for the elastic-plastic radius ρ :

$$\ln \frac{\rho}{a} + \frac{1}{3} - \frac{p_0}{2\sigma_0} + \frac{k + 4\mu/3}{k} \frac{(\alpha/\mu)}{1 - (\alpha/\mu)} \cdot \left[\frac{1}{3} \left(\frac{\rho}{a} \right)^3 - \frac{p_0}{2\sigma_0} \right] = 0 \quad (10)$$

Figure 2 shows the solutions for ρ/a as a function of $p_0/2\sigma_0$ for several values of α/μ and for a Poisson ratio equal to 1/4. In the limiting cases $\alpha/\mu = 0$ and 1 (ideal plasticity and elasticity, respectively), ρ/a is independent of the Poisson ratio, as is evident from (10). Note that for a Poisson ratio of 1/4 the factor $(k + 4\mu/3)/k$ is equal to 9/5. All curves start with the same slope at the critical pressure, the pressure at which $\rho = a$, which is equal to $2\sigma_0/3$. Below this pressure there is no plasticity.

The quantities of interest can be obtained from the value of ρ/a and the following formulas:

$$\sigma = (\rho/r)^3 \sigma_0 \quad (r \geq \rho) \quad (11a)$$

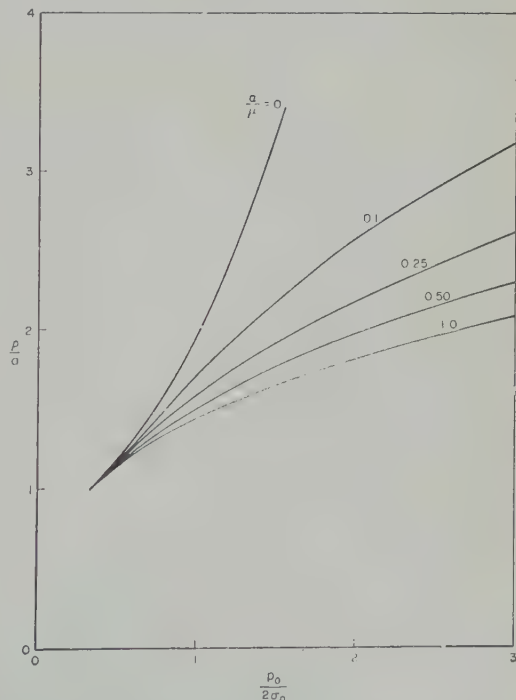


Fig. 2. The elastic-plastic radius as a function of $p_0/2\sigma_0$ for various values of α/μ and a Poisson ratio of 1/4.

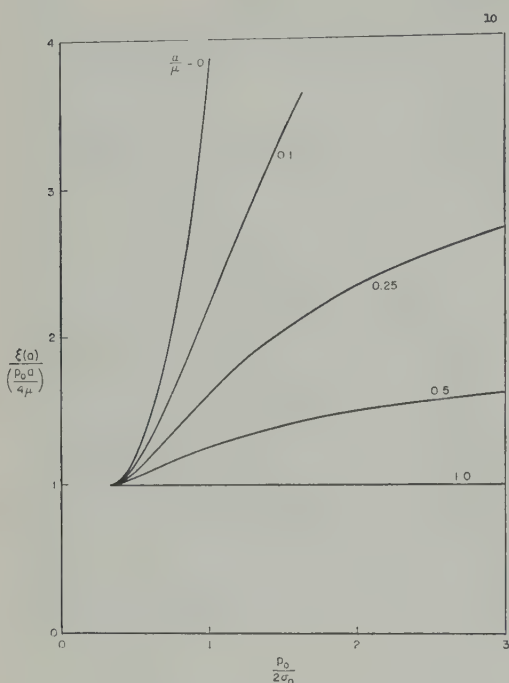


Fig. 3. The (inward) displacement at the cavity wall ($r = a$) as a function of $p_0/2\sigma_0$ for various values of α/μ and a Poisson ratio of 1/4. Note that the displacement is given as a ratio to the elastic value $p_0 a / 4\mu$.

$$\sigma = \frac{k}{k + 4\alpha/3} \frac{\alpha}{\mu} \sigma_0 \left[\frac{k + 4\mu/3}{k} \left(\frac{\rho}{r} \right)^3 + \frac{\mu}{\alpha} - 1 \right] \quad (r \leq \rho) \quad (11b)$$

$$\sigma_r = -p_0 + 2\sigma/3 \quad (r \geq \rho) \quad (12a)$$

$$\sigma_r = -p_0 + 2\sigma/3 + \frac{2k}{k + 4\alpha/3} \cdot \left(1 - \frac{\alpha}{\mu} \right) \sigma_0 \ln \frac{\rho}{r} \quad (r \leq \rho) \quad (12b)$$

$$\xi = -\frac{\sigma_0 \rho}{6\mu} \left(\frac{\rho}{r} \right)^2 \quad (r \geq \rho) \quad (13a)$$

$$\xi = -\frac{k + 4\mu/3}{k + 4\alpha/3} \frac{\sigma_0 \rho}{6\mu} \left(\frac{\rho}{r} \right)^2 + \frac{2}{3} \left(1 - \frac{\alpha}{\mu} \right) \cdot \frac{\sigma_0}{k + 4\alpha/3} r \left(\ln \frac{\rho}{r} + \frac{1}{3} \right) \quad (r \leq \rho) \quad (13b)$$

Of particular interest are the displacement $\xi(a)$ and the shear stress $\sigma(a)$ at the cavity wall

$r = a$. These quantities are shown in Figures 3 and 4 as functions of $p_0/2\sigma_0$ for various values of α/μ , again for a Poisson ratio of 1/4. In these figures the unit of displacement is the elastic value $p_0 a / 4\mu$, and the unit of shear stress is the elastic value $3p_0/2$.

After the explosion. The explosion subjects the cavity wall to a pressure-time history which can be calculated quite accurately from the equations of hydrodynamics by assuming the cavity is a rigid spherical chamber. Typical of the pressure as a function of time appears shown in Figure 5 with a very short duration spike caused by the reflected shock wave settling rapidly to a steady pressure p_∞ .

If it were necessary to solve the time-dependent elastoplastic equations in order to find the seismic signal corresponding to this pressure-time history, we would be faced with a very difficult mathematical problem. Moreover, we do not have the physical data needed to include the effects of strain rate, which are probably important in the transients. Fortunately, the seismic signal at a great distance from the explosion under

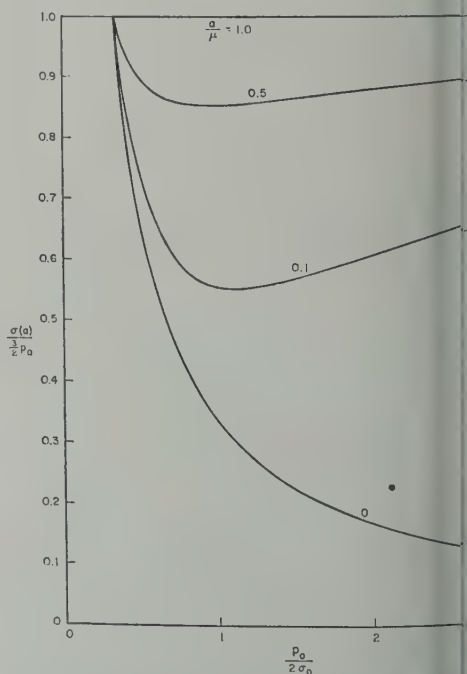


Fig. 4. The shear stress at $r = a$ as a function of $p_0/2\sigma_0$ for various values of α/μ and a Poisson ratio of 1/4.

filtering action of the earth, which allows low-frequency components of the signal to be transmitted. According to the theory of *r*, *Martinelli, and Teller* [1959], the amplitude of a low-frequency signal is simply proportional to the permanent volume expansion of a spherical surface which completely encloses the region of inelastic behavior. It follows that to obtain a good approximation of the distant seismic signal all we need to do is solve the elastoplastic equations corresponding to a zero pressure in the cavity, specifically the case p_{∞} (Fig. 5).

The approximation here amounts to ignoring strain hardening caused by the transients. Because of the nonlinearity of plasticity, this hardening can affect the static solution and the distant signal. The magnitude of this has not been carefully investigated, but we believe it to be small, and it will be ignored here. For our purposes we will start with stresses and strains existing around the cavity as calculated in the previous section and then turn on the pressure p_{∞} adiabatically and monotonically, so that each point in the medium moves in the stress-strain diagram along paths similar to the dashed paths in Fig. 1.

Evidently, if p_{∞} is small enough, the point *B* in Figure 1 will not be exceeded anywhere in the medium, in which case the entire medium deforms elastically, and the plasticity which occurred during the making of the cavity is irrelevant to the distant seismic signal. On the other hand, the stress and strain distribution around the cavity is affected by the plastic deformation and can be obtained by adding to the elastic solutions of the preceding section the appropriate elastic solutions, namely

$$\sigma_r = -p(a/r)^3 \quad (14a)$$

$$\sigma = 3\sigma_r/2 \quad (14b)$$

We have now answered the first question raised in the introduction, concerning the size of the plastic zone around the cavity, showing that it has no effect whatsoever on the decoupling factor provided the pressure in the hole is small enough. This leads naturally to the second question, namely what is the maximum pressure in the hole so that the point *B* is not exceeded anywhere in the medium?

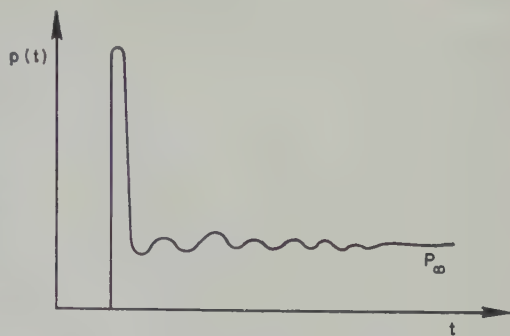


Fig. 5. Pressure vs. time on cavity wall.

As the pressure in the hole (p_{∞}) is increased, each point of the medium moves toward its point *B* on the stress-strain diagram, but clearly the points with radius $r = a$ get there first; we denote the value of p_{∞} at which this occurs by p_{p1} . From Figure 1 and equation 14b we see that

$$-(3/2)p_{p1} + \sigma(a) = -\sigma(a) \quad (15)$$

and hence, referring to (11b) for $\sigma(a)$,

$$\frac{p_{p1}}{2\sigma_0} = \frac{2}{3} \left[\left(\frac{k + 4\mu/3}{k + 4\alpha/3} \right) \frac{\alpha}{\mu} \left(\frac{\rho}{a} \right)^3 + \frac{k}{k + 4\alpha/3} \left(1 - \frac{\alpha}{\mu} \right) \right] \quad (16)$$

The quantity ρ/a is a function of $p_0/2\sigma_0$ through (10) (cf. Fig. 2). Values of $p_{p1}/2\sigma_0$ are plotted in Figure 6 as a function of $p_0/2\sigma_0$ for various values of α/μ and a Poisson ratio of 1/4. Referring to the figure, we see that if $p_0 < 4\sigma_0/3$, p_{p1} is greater than p_0 for all values of α/μ . If α/μ is greater than about 0.1, p_{p1} is greater than p_0 for any p_0 .

To understand the significance of these results we must know the stress-strain diagram quantitatively. Since salt is a medium of particular interest for decoupling cavities, we have plotted its stress-strain diagram as given by *Handin* [1953] in Figure 7. Handin's data are actually for many different cases of triaxial loading, and we have chosen the curve that seems to us most appropriate to our problem. Handin's data also are not sufficient for the determination of the elastic portion of the stress-strain curve and for this purpose we have used some measurements from the Bureau of Mines [*Nicholes, Hooker, and Duwall*, 1960], specifically a value $\mu \approx 100$ kilobars, and a Poisson ratio of 1/4. A reasonable two-segment fit to the curve in Figure 7, assuming

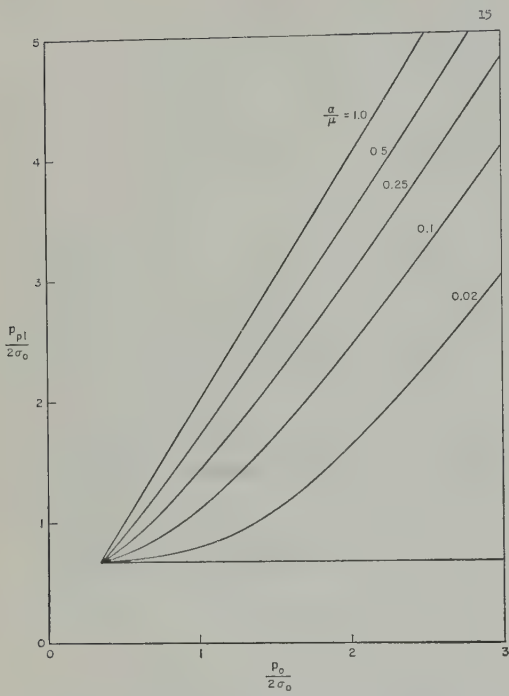


Fig. 6. Pressure (p_{p1}) in the cavity at which the medium stops responding elastically to the explosion pressure, as a function of $p_0/2\sigma_0$ for various values of α/μ and a Poisson ratio of 1/4.

that σ 's of interest do not exceed a few hundred bars, is $\sigma_0 \approx 150$ bars and $\alpha/\mu \approx 0.1$.

With these values of σ_0 and α/μ it follows from the results in the last paragraph that plasticity sets a less stringent limit on the cavity volume than the requirement that the average pressure in the cavity be less than p_0 (*a fortiori* if the average pressure is less than $p_0/2$). Therefore the answer to the second question in the introduction is that for fully decoupled explosions, plasticity is unimportant at any depth.

Overdriven holes. We now come to the last question, concerning the effect of plasticity on decoupling with overdriven holes. As was mentioned in the introduction, plasticity did not play an important role in the Cowboy experiments because of their relatively shallow depth. However for the same degree of overdriving, i.e., the same value of p_∞/p_0 , for cavities at a depth of ~ 1 km, plasticity cannot be ignored. In this case the reduction in the decoupling factor may be due to plasticity rather than to cracking.

That plasticity can lead to a big reduction in decoupling can be seen from the following simple

argument. Because of plasticity the value of stress differences $|\sigma_r - \sigma_t|$ is limited to something in the neighborhood of the yield stress σ_0 . If p_∞ is large compared to σ_0 , as it will be for overdriven shots at great depth, it follows that σ_r will be approximately equal to σ_t , which is the characteristic property of a liquid. For a liquid, (static) pressures are transmitted virtually without attenuation, which implies large displacements.

Of course a plastic medium has more strength than a liquid, but how much more depends on the detailed stress-strain diagram or, in a two-segment fit, on the values of α/μ and β . To obtain a value for the reduction in decoupling factor due to plasticity, we have gone through an analysis similar to that of the preceding section but with p_∞ greater than p_0 . The algebra is complicated, involving bookkeeping on the zones in which plasticity occurred once, twice, or not at all, and we merely quote the results here. The details are given in a Rand Corporation report [Mathéws, 1961].

Table 1 gives the factor by which decoupling is reduced for cavities at a depth of 1 km in salt if we assume several different fits to Handin's [1959] stress-strain diagram given in Fig. 7). As is evident in Figure 7, the best value of α/μ depends on the value of p_∞ and hence on the degree of overdriving. For a small amount of overdriving a reasonable value of α/μ is ~ 0.1 . For the greatest overdriving a better value of α/μ is ~ 0.02 . The extreme plastic behavior, $\alpha/\mu = 0$, has also been included for comparison. Finally, the table contains the reduction factor in decoupling as experimentally determined for the overdriven Cowboy shots.

In interpreting the data of Table 1 it is important to inquire whether in cases 1 to 3 the medium has at any time gone into tension,

TABLE 1. Reduction Factor f in Decoupling for Overdriven Cavities at a Depth of 1 km ($p_0 = 200$ bars) with Various Stress-Strain Relations Approximating That of Salt

Case	p_∞/p_0	
	1	2.5
1. $\sigma_0 = 150, \alpha/\mu = 0.1$	~ 1	1.4
2. $\sigma_0 = 150, \alpha/\mu = 0.02$	~ 1	1.8
3. $\sigma_0 = 150, \alpha/\mu = 0$ (ideal)	~ 1	2.0
4. Cowboy ($p = 50$ bars)	~ 1	~ 4

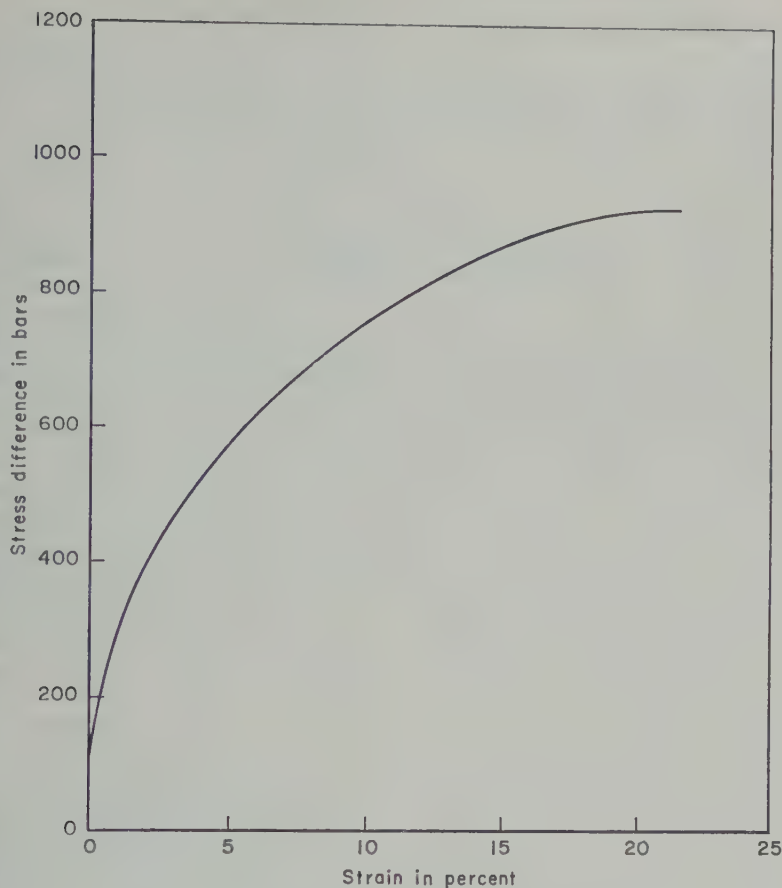


Fig. 7 Stress-strain curve for rock salt (after Handin).

0, since in that event cracking as well as plasticity would have to be considered. A detailed examination of the plastic solutions shows that the medium does not go into tension for values of p_{∞}/p_0 between 1 and 5, and therefore cracking due to tension does not occur. We note in passing that in the case of ideal plasticity, $\alpha/\mu = 0$, the medium never goes into tension no matter how large p_{∞}/p_0 becomes; in case 2, $\alpha/\mu = 0.02$, tension occurs when $p_{\infty}/p_0 = 1.2$; and in case 1, $\alpha/\mu = 0.1$, tension occurs when $p_{\infty}/p_0 = 7.5$.

Besides cracking due to tension, there is the possibility previously mentioned of crack propagation due to the explosion gases penetrating pre-existing cracks, which sets the limit for the behavior at $p_{\infty}/p_0 \sim 1$. If this mechanism is the controlling form of inelasticity in Cowboy experiments, if the mechanism continues to operate for driven shots at a depth of ~ 1 km, then the

appropriate change in decoupling factor is probably close to that observed in the Cowboy experiments. If, however, the Cowboy inelasticity was due to something else, for instance the large transient pressure, or if the occurrence of plasticity has a sealing effect on the pre-existing cracks, which seems possible, then the appropriate change in decoupling factor is determined by plasticity. In this case Table 1 shows that the change in decoupling factor for deep overdriven cavities may be even smaller than that observed in the Cowboy experiments.

Cowboy inelasticity. We have said a number of times that the reduction in decoupling observed in the Cowboy experiments was probably due not to plasticity but to cracking, either because the medium was subjected to tension or because gases leaked into pre-existing cracks. In this section we shall support this statement by showing that the actual observed values for the

change in the decoupling factors are consistent with this picture.

Making use of the theory of Latter, Martinelli and Teller [1959], we can write the factor by which the decoupling is reduced as

$$f = \Delta p r_{ei}^3 / p_{\infty} a^3 \quad (17)$$

where r_{ei} is the smallest radius outside which the medium is elastic, Δp is the increase in pressure at the radius r_{ei} , p_{∞} is the average pressure in the cavity, and a is the radius of the cavity, as previously defined. The quantity Δp —the maximum radial stress which the medium can sustain without becoming inelastic—for rock-like materials which have no strength in tension must be something in the neighborhood of the overburden pressure p_0 , i.e.,

$$\Delta p = k p_0 \quad (18)$$

where $k \sim 1$. To relate r_{ei} to Δp and p_{∞} we must know how the radial stress depends on the radius. If we assume that the pressure varies inversely with the n th power of the radius, the connection for a given value of p_{∞} is

$$\Delta p = p_{\infty} (a/r_{ei})^n \quad (19)$$

where the exponent n may itself depend on p_{∞} . From (17), (18), and (19)

$$f = (p_{\infty}/k p_0)^{3/n-1} \quad (20)$$

By inserting the values of f from the Cowboy shots (Table 1) we obtain an equation for each of the two cases in which $p_{\infty}/p_0 = 2.5$ and 5. These two equations are not sufficient to determine uniquely the value of k and the two values of n . However, for a reasonable value of k , in the range $p_0/2$ to $3p_0/2$, it turns out that the two values of n are roughly equal to each other and lie in the range 1 to 1.5.

What does such a value of n , between 1 and 1.5, imply? From (1), assuming the power-law dependence for the radial stress, we see that the ratio of the hoop stress to the radial stress is simply

$$\sigma_i/\sigma_r = 1 - n/2 \quad (21)$$

Now there are two well-known cases in which the power-law dependence of the radial stress on distance is exact. One of these is elasticity, for which $n = 3$ and therefore $\sigma_i/\sigma_r = -1/2$. At the

other extreme is (liquid) hydrostatics for which $n = 0$ and $\sigma_i/\sigma_r = 1$. Between these two cases of an elastic solid and a liquid ($n = 3$ and $n = 0$) one can imagine all kinds of media with varying degrees of strength leading to values of n between 0 and 3. In particular, if $n = 2$, $\sigma_i/\sigma_r = 0$, the hoop stress is zero, which in some ideal sense describes a medium with open cracks such as might be constructed with building blocks.

We are now in a position to interpret a value of n which lies in the range of 1 to 1.5. Such a value of n implies that σ_i/σ_r is between $1/3$ and $1/2$ and corresponds to a medium which is intermediate between the building-blocks model and a liquid.

(After writing this paper, we learned that Dr. Norman Haskell has investigated the effect of inelasticity based on the Coulomb-Mohr yield condition. In Haskell's work the power-law dependence of pressure on distance, our equation 19—which we merely assume—is actually derived. Dr. Haskell's paper can be found elsewhere in this issue of the *Journal of Geophysical Research*.)

REFERENCES

- Handin, John, Application of high pressure in geophysics: Experimental rock deformation, *Trans. Am. Soc. Mech. Engrs.*, **75**, 315, 1953.
- Herbst, R. F., G. C. Werth, D. L. Springer, Underground large cavities to reduce seismic waves from underground explosions, *J. Geophys. Research*, **66**, 1961.
- Latter, A. L., *Verbatim Record of Technical Working Group 2, Conference on the Discontinuities of Nuclear Weapon Tests*, GEN/DNT/TWG.2, December 2, 1959.
- Latter, A. L., R. E. LeLevier, E. A. Martinelli, and W. G. McMillan, A method of concealing underground nuclear explosions, *J. Geophysical Research*, **66**, 943, 1961.
- Latter, A. L., E. A. Martinelli, and E. Teller, Dynamic scaling law for underground explosions, *J. Fluids*, **2**, 280, 1959.
- Mathews, J., Static deformation of a plastic medium, *Rand Corp., Research Mem. RM-2686*, January 18, 1961.
- Murphey, B. F., Particle motions near explosion in halite, *J. Geophys. Research*, **66**, 947, 1961.
- Nicholes, H. R., V. Hooker, and W. I. Doolittle, Dynamic rock mechanics investigations Part I, Cowboy, *Rept. ARRL 38-3.2*, U. S. Bureau of Mines Applied Physics Research Laboratory, College Park, Md., Sept. 1, 1960.

(Manuscript received June 10, 1961.)

A Static Theory of the Seismic Coupling of a Contained Underground Explosion

NORMAN A. HASKELL

*Geophysics Research Directorate
Air Force Cambridge Research Laboratories
Bedford, Massachusetts*

Abstract. According to the theory of Latter, Martinelli, and Teller the amplitude of the distant seismic signal from a completely contained underground explosion is determined by the permanent displacement produced in the neighborhood of the source. A static-equilibrium theory of this displacement is developed. A Coulomb-Mohr type of yield condition is used to determine the stresses in the near zone where the stresses are beyond the elastic limit. If the internal friction parameter that occurs in the Coulomb-Mohr yield condition is treated as a phenomenological constant, to be determined by the seismic data, it is possible to obtain reasonably good agreement with the relative amplitudes of the seismic signals observed in the Project Cowboy series of chemical explosions in cavities of various sizes in salt. The indicated value of the friction parameter is, however, appreciably less than the values usually observed in compression tests on unconsolidated materials. The theory is also consistent with the observed size of the cavity produced in tuff by the underground nuclear explosion Rainier, but an even smaller value of the friction parameter must be assumed in this case.

Introduction. According to the theory of Latter, Martinelli, and Teller [1959], the amplitude of the seismic waves recorded at large distances from a completely contained underground explosion is determined by the permanent displacement produced in the neighborhood of the source. A conclusion follows from the fact that high frequencies are attenuated by the earth more rapidly than low frequencies. The high-frequency components generated by the impact of the shock wave on the walls of the explosion cavity therefore make only a relatively small contribution to the distant signal, most of which is attributed to the quasi-permanent increase in cavity pressure produced by the explosion. Since the final state is one of static equilibrium, these considerations suggest that a parameter proportional to the amplitude of the distant seismic signal may be derived from a static theory without knowledge of the details of the dynamic processes by which the final state is reached.

At this point it may be advisable to insert a clarification of the expressions 'permanent displacement' and 'final state.' In any real case the explosion products must eventually leak away as well as cool down to ambient temperature, and the cavity pressure will drop. However, if

the explosion is to be described as 'contained,' these processes must be slow on the time scale of the explosion itself. Thus, before its ultimate decay, the pressure in the cavity must level off to a more or less constant plateau, and it is this state that we shall treat as the final one. This means that when the distant seismic signal is described as being dominantly of low frequency we do not mean that the corresponding periods are long compared with the diffusive decay time of the cavity pressure but only that the corresponding wavelengths are long compared with the dimensions of the cavity.

Coulomb-Mohr yield criterion. In the final state the mechanical properties of the material immediately surrounding the explosion cavity will not be the same as those of the original undisturbed medium, since this material will have been previously subjected to a large transient stress system greatly exceeding its fracture strength. It will therefore be assumed that the cavity is surrounded by a crushed zone and that within this zone the material has the mechanical properties of a nearly incoherent granular aggregate. Experimental investigations in soil mechanics [Terzaghi and Peck, 1948] have shown that the maximum stress difference in many such materials is fairly well described by the Coulomb-

Mohr yield criterion. For our purposes this may be stated in the following form. Let σ_1 , σ_2 , and σ_3 be the principal stresses at a point in the medium, with $\sigma_3 > \sigma_2 > \sigma_1$ (tension taken as positive). Then the maximum stress difference at the yield point is given by

$$\sigma_3 - \sigma_1 \leq (2\sigma_0 - \sigma_1 - \sigma_3) \sin \phi \quad (1)$$

The tangent of the angle ϕ is the Coulomb coefficient of friction on internal fracture surfaces. The quantity σ_0 is related to the uniaxial tensile strength of the material, σ_T , by the relation

$$2\sigma_0 = \sigma_T(1 + \sin \phi)/\sin \phi \quad (2)$$

In the numerical examples to be discussed later we shall assume a negligible tensile strength and set $\sigma_0 = 0$; however, for the sake of generality we shall not make this assumption in the formal development.

Hubbert and Rubey [1959] have considered the effect of a fluid filling on the mechanics of a porous solid and have derived a modification of the Coulomb-Mohr yield condition which is simply equivalent to replacing the quantity σ_0 in (1) by $\sigma_0 - p \tan \phi$, where p is the pressure in the pore fluid. The limiting condition on the difference between the greatest and least principal stresses then becomes

$$\sigma_3 - \sigma_1 \leq [2(\sigma_0 - p \tan \phi) - (\sigma_1 + \sigma_3)] \sin \phi \quad (3)$$

Static solution for a spherically symmetrical explosion. Since the effective strength of the crushed material depends upon the magnitude of the principal stresses, the part of these stresses due to gravity cannot be ignored. However, to preserve the simplicity associated with spherical symmetry, we shall not treat the effect of gravity directly. Instead we shall approximate its effect by considering the medium to be subjected to a constant hydrostatic pressure, P_0 , equal to the overburden pressure at the depth of the shot. Then in a spherical coordinate system (r , θ , φ) with origin at the center of the shot cavity the only equation of equilibrium that is not satisfied identically by the assumption of spherical symmetry is the radial equation

$$dT_{rr}/dr + 2(T_{rr} - T_{\theta\theta})/r = 0 \quad (4)$$

where T_{rr} and $T_{\theta\theta}(=T_{\varphi\varphi})$ are the only non-vanishing components of the stress tensor.

We assume that the explosion takes place in a

spherical cavity of initial radius r_0 , which expands to a radius r_1 in the final state, and that beyond some radius $r_2 > r_1$ the medium behaves in accordance with linear elastic theory. Within the explosion cavity the pressure is assumed to be given by an ideal gas law with a quantity of energy W introduced into the initial volume followed by adiabatic expansion to the final cavity volume. In the region between r_1 and r_2 it is assumed that the difference between the principal stresses has everywhere the maximum value permitted by the yield condition (3), which $\sigma_3 = T_{\theta\theta}$ and $\sigma_1 = T_{rr}$. The combination of (4) with (3), interpreted as an equilibrium condition, permits a determinate solution for the stresses. The radial displacement $U(r)$ is determined in this region from the conservation of mass, under the assumption that the bulk compressibility of the material is the same as that in the elastic region. In the region outside r_2 the stresses are superimposed on the initial uniform hydrostatic pressure P_0 and are connected with the displacements by the usual elastic relationships [Love, 1926], so that

$$T_{rr} = -P_0 + (\lambda + 2\mu) dU/dr + 2\lambda U/r$$

$$T_{\theta\theta} = -P_0 + \lambda dU/dr + 2(\lambda + \mu)U/r$$

The boundary conditions at the radius r_1 are

$$T_{rr}(r_1) = -P$$

where P is the final cavity pressure, and

$$U(r_1) = r_1 - r_0$$

At $r = r_2$ the required boundary conditions are continuity of $T_{rr}(r_2)$ and $U(r_2)$. In addition we shall impose the condition that $T_{\theta\theta}(r_2)$ be continuous. The latter would not be a required condition if the surface $r = r_2$ were a boundary separating two different elastic media; in that case r_2 would have been assigned a specific value as one of the conditions of the problem, it would not generally be possible to find a solution that would be compatible with continuity of $T_{\theta\theta}$. In the present case, however, the boundary is not one between different media but between different modes of behavior of the same medium, and r_2 is one of the unknowns of the problem. The value of r_2 is, in fact, determined by the condition that at r_2 frictional forces are insufficient to prevent relative movement on slip planes, and the equality sign holds in (3); outside r_2 frictional forces are great enough to prevent internal

age, and the inequality sign holds. It seems possible to assume that the transition takes place without abrupt discontinuity in the stress difference $T_{\theta\theta} - T_{rr}$, and therefore, since T_{rr} is continuous, $T_{\theta\theta}$ is also. However, it should be emphasized that this is introduced as a hypothesis and not as a statement derivable *a priori* from more basic principles.

Under the ideal gas assumption the energy of explosion is

$$W = 4\pi r_0^3 P_i / 3(\gamma - 1) \quad (9)$$

where P_i is the initial pressure in the cavity of radius r_0 , and γ is the specific heat ratio of the gas. If the expansion from the radius r_1 is adiabatic, the final cavity pressure is

$$P_i (r_0/r_1)^{3\gamma} = \frac{3(\gamma - 1)W}{4\pi r_0^3} \left(\frac{r_0}{r_1}\right)^{3\gamma} \quad (10)$$

In the region between r_1 and r_2 we have from (3)

$$T_{rr} = [2(\sigma_0 - p \tan \phi) - (T_{\theta\theta} + T_{rr})] \sin \phi$$

which may also be written in the form

$$T_{\theta\theta} = 2k(T_{rr} - \sigma_1)/(1 + k) \quad (11)$$

where we have set $k = \sin \phi$ and $\sigma_1 = \sigma_0 - p$. Substituting (11) into (4) gives a simple equation for T_{rr} ,

$$\frac{dT_{rr}}{dr} + \frac{4k}{r(1+k)} (T_{rr} - \sigma_1) = 0 \quad (12)$$

The solution is

$$T_{rr} = \sigma_1 - A r^{-m} \quad (13)$$

where $m = 4k/(1+k)$. The boundary condition determines the constant of integration A ,

$$T_{rr} = \sigma_1 - (\sigma_1 + P) \left(\frac{r_1}{r}\right)^m \quad (14)$$

In (11) we then have

$$T_{\theta\theta} = \sigma_1 - \frac{1-k}{1+k} (\sigma_1 + P) \left(\frac{r_1}{r}\right)^m \quad (15)$$

To determine the displacement in the region between r_1 and r_2 we let a be the initial radial coordinate of a particle whose final coordinate is r . Under the assumption of spherically symmetrical displacement the dilatation Δ is

$$\Delta = \frac{r^2 dr - a^2 da}{a^2 da} = \frac{r^2}{a^2} \frac{dr}{da} - 1$$

The displacement, considered as a function of the final coordinate r , is

$$U(r) = r - a(r) \quad (16)$$

and, considering Δ as a function of r , we may write

$$a^3 = (r - U)^3 = 3 \int_{r_1}^r \frac{r^2 dr}{1 + \Delta} + \text{const.}$$

The constant of integration is determined from the boundary condition (8) to be r_0^3 ; hence we have

$$U = r - \left[r_0^3 + 3 \int_{r_1}^r \frac{r^2 dr}{1 + \Delta} \right]^{1/3} \quad (17)$$

Since $\Delta \ll 1$, we replace $(1 + \Delta)^{-1}$ in the integral by $(1 - \Delta)$ and integrate the first term, giving

$$U = r - \left[r_0^3 + r^3 - r_1^3 - 3 \int_{r_1}^r \Delta r^2 dr \right]^{1/3} \quad (18)$$

The dilatation and displacement in this expression are measured from the initial state, which, according to our assumption, is one of uniform pressure P_0 . In expressing Δ in terms of the principal stresses we must therefore subtract $-P_0$ from each of the total stress components, i.e.

$$\Delta = \frac{1}{3\lambda + 2\mu} (T_{rr} + 2T_{\theta\theta} + 3P_0) \quad (19)$$

where λ and μ are the Lamé elastic constants of the medium. With the values of T_{rr} and $T_{\theta\theta}$ given by (14) and (15), this becomes

$$\Delta = \frac{1}{3\lambda + 2\mu} \left[3(\sigma_1 + P_0) - (\sigma_1 + P) \left(\frac{3-k}{1+k} \right) \left(\frac{r_1}{r} \right)^m \right]$$

From the definition (13) of m we have $(3-k)/(1+k) = 3-m$, and the integral in (18) becomes

$$\int_{r_1}^r \Delta r^2 dr = \frac{r_1^3}{3\lambda + 2\mu} \left[\{\sigma_1 + P_0\} \left\{ \left(\frac{r}{r_1} \right)^3 - 1 \right\} - \{\sigma_1 + P\} \left\{ \left(\frac{r}{r_1} \right)^{3-m} - 1 \right\} \right] \quad (20)$$

This, with (18), completes the determination of U in the region $r_1 < r < r_2$. Actually, we are interested in the value of U only at $r = r_2$, and, since r_2^3 is, in all practical cases, much larger than the absolute value of the rest of the bracketed expression in (18), we may expand the bracket, retaining only first-order terms, and obtain

$$r_2^2 U(r_2) = \frac{r_1^3 - r_0^3}{3} + \frac{r_1^3}{3\lambda + 2\mu} \left[\{\sigma_1 + P_0\} \left\{ \left(\frac{r_2}{r_1} \right)^3 - 1 \right\} - \{\sigma_1 + P\} \left\{ \left(\frac{r_2}{r_1} \right)^{3-m} - 1 \right\} \right] \quad (21)$$

The solution of (4), (5), and (6) in the elastic region $r > r_2$ is well known [Love, 1944] and can be written immediately as

$$U = Br^{-2} \quad (22)$$

$$T_{rr} = -P_0 - 4\mu Br^{-3} \quad (23)$$

$$T_{\theta\theta} = -P_0 + 2\mu Br^{-3} \quad (24)$$

Continuity of U at r_2 shows that the constant $B = r_2^2 U(r_2)$ which is given by (21). Continuity of T_{rr} and $T_{\theta\theta}$ gives the relations

$$(\sigma_1 + P_0) - (\sigma_1 + P)(r_1/r_2)^m = -4\mu Br_2^{-3} \quad (25)$$

and

$$(\sigma_1 + P_0) - \left(\frac{1-k}{1+k} \right) (\sigma_1 + P) \left(\frac{r_1}{r_2} \right)^m = 2\mu Br_2^{-3} \quad (26)$$

Since P is a known function of r_1 and r_0 (equation 10), equations 21, 25, and 26 determine the three unknowns B , r_1 , and r_2 when r_0 and the other constants of the problem are given. The constant B is the parameter which, according to the theory of Latter, Martinelli, and Teller, is proportional to the amplitude of the distant seismic signal and thus provides a measure of the seismic coupling of the explosion.

For the purpose of numerical computations it is convenient to transform these equations as follows:

Elimination of B between (25) and (26) gives the relation

$$\sigma_1 + P = \frac{3(1+k)}{3-k} (\sigma_1 + P_0) \left(\frac{r_2}{r_1} \right)^m$$

and inserting this in (21) gives

$$B = \frac{r_1^3 - r_0^3}{3} - \frac{r_1^3 (\sigma_1 + P_0)}{(3\lambda + 2\mu)} \cdot \left[1 + \frac{4k}{3-k} \left(\frac{r_2}{r_1} \right)^3 - \frac{3(1+k)}{3-k} \left(\frac{r_2}{r_1} \right)^m \right]$$

From (26) and (27) we also have

$$B = r_2^3 k (\sigma_1 + P_0) / \mu (3 - k)$$

Equating these values of B gives an equation for $(r_0/r_1)^3$ in terms of (r_2/r_1)

$$\left(\frac{r_0}{r_1} \right)^3 = 1 - \frac{3k(\sigma_1 + P_0)}{\mu(3-k)} \left(\frac{r_2}{r_1} \right)^3 - \frac{3(\sigma_1 + P_0)}{(3\lambda + 2\mu)} \cdot \left[1 + \frac{4k}{3-k} \left(\frac{r_2}{r_1} \right)^3 - \frac{3(1+k)}{3-k} \left(\frac{r_2}{r_1} \right)^m \right]$$

To determine B , r_1 , and r_2 as functions of r_0 we may assume a series of values of r_2/r_1 and determine r_0/r_1 from (30) and P from (10). The corresponding initial radius is then computed from (10),

$$r_0^3 = \frac{3(\gamma - 1)W}{4\pi P} \left(\frac{r_0}{r_1} \right)^{3\gamma}$$

and B is computed from (29).

If the initial cavity is sufficiently large, the stresses on the cavity wall will remain within the elastic limit, and the value of B is given by the known solution of the elastic problem:

$$B = Pr_0^3 / 4\mu = 3(\gamma - 1)W / 16\pi\mu$$

The upper limit of the values of r_0 for which equation (31) may be considered valid may be taken as the point where B , as given by (31), is equal to the value given by (32). For larger values of r_0 , B has the constant value given by the latter equation.

In a chemical explosion the minimum value of r_0 is equal to the radius of the explosive charge. The extremely high temperature of a nuclear explosion will vaporize the surrounding material and then the minimum value of r_0 is to be interpreted as the radius of rock vaporization rather than that of the nuclear device itself.

Comparison with Project Cowboy data. The relative amplitudes of the seismic waves produced by high-explosive shots fired in sphere

ies and in tamped holes in a salt dome have been discussed by *Herbst, Werth, and Springer* [1]. For comparison with these data the coupling parameter B has been computed as a function of r_0 for a 2000-lb explosion at a depth of 90 ft in salt and various assumed values of the internal friction parameter k . Assumed values of the other pertinent constants are:

$$= 5.117 \times 10^7 \text{ dynes/cm}^2.$$

$$= 0.$$

$$= 1.697 \times 10^{11} \text{ dynes/cm}^2.$$

$$+ 2\mu = 8.485 \times 10^{11} \text{ dynes/cm}^2.$$

$$= 4.185 \times 10^{16} \text{ ergs.}$$

$$= 1.2.$$

These constants are calculated on the basis of a reported compressional wave velocity of 4.9 sec in salt with the assumption that Poisson's ratio is $\frac{1}{4}$ (i.e., $\lambda = \mu$) and that the density is 2.17 g/cm^3 . The results are plotted in Figure 1. The experimental data were obtained from shots in cavities of 12-ft and 30-ft diameter at depths of about 800 feet and from tamped shots in cased bore holes about 110 feet deeper. The loading density of the tamped shots was assumed that a 2000-lb charge occupied a sphere

about 4 feet in diameter. If we interpret a shot with energy yield different from that of the 2000-lb charge as a scaled model of a 2000-lb shot with r_0 scaled in proportion to the cube root of the energy ratio (which is consistent with the equations developed above), the data given by Herbst, Werth, and Springer (cf. their Fig. 10) are equivalent to 2000-lb shots in cavities of the radii listed in column 2 of Table 1. Column 3 gives the average observed decoupling factors, i.e., the ratios of the seismic signal amplitude for a fully tamped shot to the signal amplitude of a shot of the same size in the given cavity. According to the interpretation of the quantity B , as proportional to the amplitude of the distant seismic signal, this is $B(\text{tamped})/B(r_0)$. In columns 4a and 4b the decoupling factors are converted into absolute values of B for a 2000-lb shot on two different bases. Herbst, Werth, and Springer give the average permanent displacement produced by a 1000-lb tamped shot as 0.010 inch at a distance of 80 feet. This is equivalent to a value of $B(\text{tamped})$ of 0.151 m^3 for 1000 lb or 0.302 m^3 for 2000 lb. The figures in column 4a are calculated on this basis. On the other hand, the value of B for a cavity that

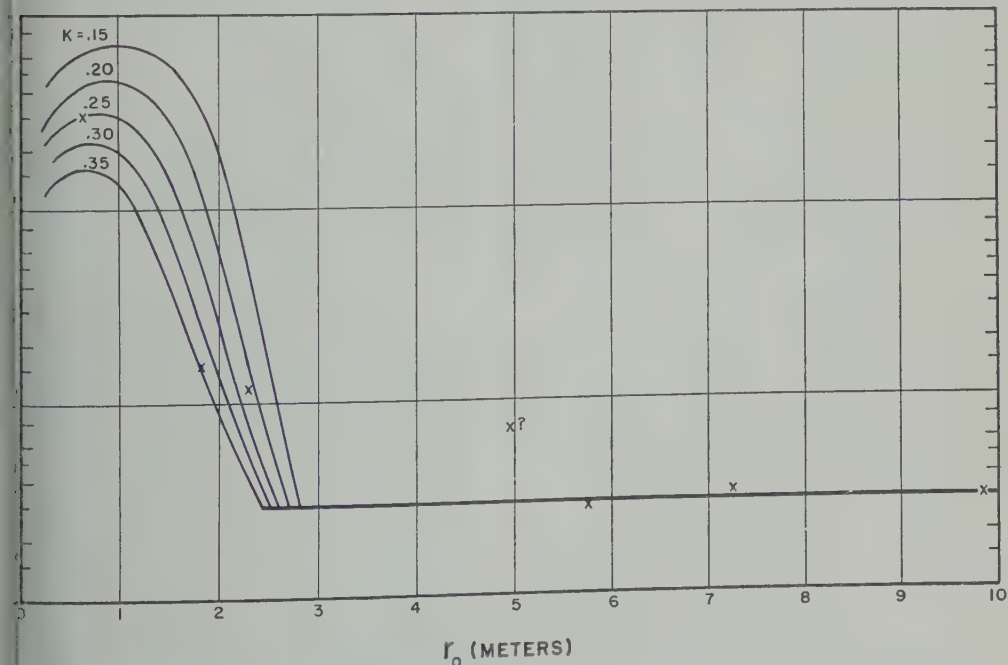


Fig. 1. B versus initial cavity radius for 2000-lb high-explosive shots in salt. Crosses indicate values scaled from Project Cowboy data.

TABLE 1. Project Cowboy Decoupling Data

(1)	(2)	(3)	(4)	
Shot	Equiva- lent r_0 for 2000-lb Shot, meters	Decou- pling Factor	B for 2000-lb Shot, cubic meters (a)	(b)
Tamped 2000-lb	0.61	1	0.302	0.294
12-ft sphere	1.83	19	.0159	.0155
1000-lb				
12-ft sphere	2.30	25	.0121	.0118
100-lb				
12-ft sphere	4.96	42?	.00719?	.00701?
1000-lb				
30-ft sphere	5.76	105	.00288	.00280
500-lb				
30-ft sphere	7.26	90	.00336	.00327
200-lb				
30-ft sphere	9.85	100	.00302	.00294

behaves purely elastically is given by equation 32 and is equal to 0.00294 m³ for a 2000-lb shot with the constants given for such a shot. The values of B given in column 4b are calculated on the assumption that the 200-lb shot in the 30-ft sphere represents purely elastic behavior. In view of the variability of the experimental data and the uncertainties in the assumed elastic constants, the two values of B may be considered to be practically identical.

The values of B given in column 4b of Table 1 are plotted against r_0 as crosses in Figure 1. Except for the anomalous point derived from the 100-lb shot in the 12-ft sphere, the experimentally derived points are in reasonably good agreement with the theoretical curves for values of k between 0.20 and 0.35. Herbst, Werth, and Springer point out that the anomalous point is based on measurements obtained at a single temporary station. The values of the internal friction parameter k indicated by this comparison are appreciably less than those obtained from compression tests on unconsolidated sands. Terzaghi and Peck [1948, p. 82] gave the values

TABLE 2. Friction Parameter, $k = \sin \phi$ for Dry Sand

	Round Grains, Uniform	Angular Grains, well graded
Loose packing	0.48	0.56
Dense packing	.57	.72

shown in Table 2 as typical. It is clear from Figure 1 that the present theory, with values of k of the order of 0.5 or greater, would not provide even a crude fit to the Project Cowboy data. The low apparent value of k may be an indication that the deformation in the inelastic region involves both plastic flow and slippage along fracture surfaces. The behavior of an ideal plastic material may be represented within the Coulomb-Mohr theory by setting $k = 0$, assuming a finite value for the tensile strength. Thus, if both flow and fracture are involved, the effective value of k might be expected to be less than the value that would apply in the case of fracture slippage alone.

Application to Rainier shot. No experimental data have been obtained on the effect of inner cavity size on the amplitude of the seismic signals from underground nuclear explosions. However, it may be of some interest to compare the present theory with the data that are available from the Rainier shot, which was an explosion of 1.7-kiloton energy yield fired in porous, water-saturated tuff at a depth of 100 feet (measured perpendicular to nearest surface).

The cavity in which Rainier was fired had a volume of about 7.1 m³. According to calculations by Nuckolls, reported by Johnson, Higgins and Violet [1959], the tuff was vaporized to a radius of 2.3 m. With a density of 2.0 g/cm³, the mass of vaporized rock was then about 8.8 × 10⁴ kg. Since the mass of material in the original shot chamber was about 10³ kg, most of the cavity gas was composed of vaporized rock (including about 24 per cent water vapor). If the ideal-gas assumption with a constant value of γ is retained, in spite of the fact that it is probably a crude approximation for such material, it is not clear what value of γ should be assumed as the best compromise over the range of pressure and temperatures under consideration. In the calculations we shall assume $\gamma = 1.4$ as in the case of a chemical explosion, without implication that this is a 'best' value to use for underground nuclear explosions.

The values of the elastic constants have been calculated from the compressional and shear wave velocities and the density given by Johnson, Higgins, and Violet. The elastic constants and the other constants used in the computations are:

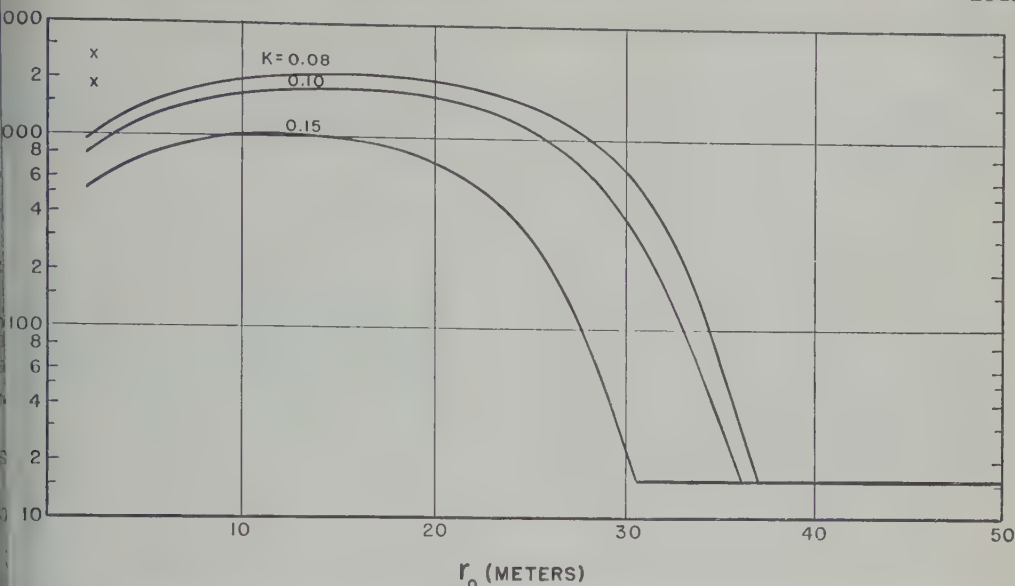


Fig. 2. B versus initial cavity radius for 1.7 kt in tuff. Crosses indicate values from Rainier shot (probably not final displacements).

$$= 4.724 \times 10^7 \text{ dynes/cm}^2.$$

$$= -2.362 \times 10^7 \tan \phi \text{ dynes/cm}^2$$

$$(\phi = \sin^{-1} k).$$

$$= 0.538 \times 10^{11} \text{ dynes/cm}^2.$$

$$+ 2\mu = 1.361 \times 10^{11} \text{ dynes/cm}^2.$$

$$= 7.11 \times 10^{19} \text{ ergs.}$$

$$= 1.2.$$

the tuff was water saturated, the Hubbert term in equation 3 for the effect of pore pressure is included, with the assumption of saturation extended to the surface. In view of the aridity of the site this is undoubtedly not true, but in the present case the effect of pore pressure term is small.

Computed values of B as a function of r_0 are shown in Figure 2 for $k = 0.08, 0.10$, and 0.15 . Values of r_1 versus r_0 are plotted in Figure 3.

Hubbert [1958] gave values of the displacement obtained by double integration of accelerometer readings at two points at distances of 371 ft (113 m) and 451 ft (137 m) vertically above the shot. The displacements measured at the calculated time of arrival of the wave from the free surface (about 230 and 250 msec, respectively) were 14.6 and 13.7 cm, respectively. If we assume that these points are in the elastic region and that the given displacements represent what the final displacements would have been in the absence of free-

surface effects (both doubtful assumptions) the corresponding values of $B = Ur^2$ are 1.87×10^8 and $2.58 \times 10^8 \text{ m}^3$. These values are shown as crosses in Figure 2, plotted at $r_0 = 2.3 \text{ m}$, as suggested by Nuckolls' calculated radius of vaporization. In view of the uncertainty in the measured displacements and in their interpretation, an order-of-magnitude agreement with

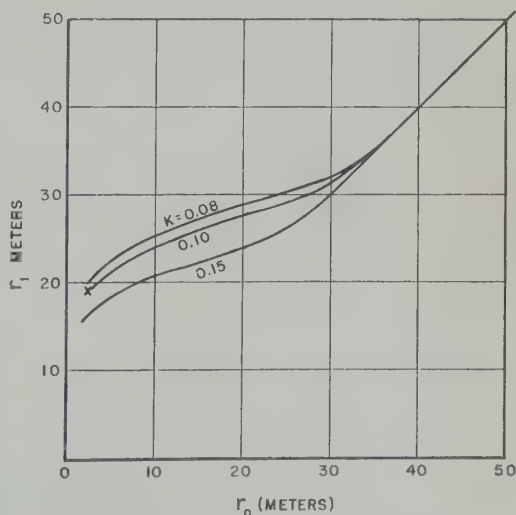


Fig. 3. Final cavity radius versus initial cavity radius for 1.7 kt in tuff. Cross indicates observed value for Rainier shot.

the calculated curves is perhaps all that we can expect.

A more convincing reason for choosing such low values of k to represent the Rainier medium is found in the final cavity radius as determined by postshot drilling of the collapsed cavity. This is given by Johnson, Higgins, and Violet as 62 ft (18.9 m). It may be seen from Figure 3 that this would fit the calculated curves of r_1 versus r_0 at $r_0 = 2.3$ m for a value of k of about 0.10.

Acknowledgement. The research reported here was carried out as part of the Geophysics Research Directorate's participation in Project Vela-Uniform of the Advanced Research Projects Agency (ARPA) of the Department of Defense. Permission by ARPA for its publication is gratefully acknowledged.

REFERENCES

Herbst, R. F., G. C. Werth, and D. L. Springer,

Use of large cavities to reduce seismic waves from underground explosions, *J. Geophys. Research*, 66, 959-978, 1961.

Hubbert, M. K., and W. W. Rubey, Mechanics of fluid-filled porous solids and its application to overthrust faulting, *Bull. Geol. Soc. Am.* 70, 1, 1959.

Johnson, G. W., G. H. Higgins, and C. E. Violet, Underground nuclear detonations, *J. Geophys. Research*, 64, 1457-1470, 1959.

Latter, A. L., E. A. Martinelli, and E. Teller, Seismic scaling law for underground explosions, *Physics of Fluids*, 2, 280, 1959.

Love, A. E. H., *Mathematical Theory of Elasticity*, 4th ed., p. 142, Dover, New York, 1944.

Perrett, W. R., Subsurface motion from a confined underground detonation, *Sandia Corp. Rept. 112*, 1529, Nov. 1958.

Terzaghi, K., and R. B. Peck, *Soil Mechanics in Engineering Practice*, John Wiley & Sons, New York, 1948.

(Manuscript received April 18, 1961.)

The Seismic Head Pulse, Reflection and Pseudo-Reflection Pulses¹

C. H. DIX

California Institute of Technology

Pasadena, California

Abstract. The shapes of the seismic head pulse, the ordinary reflection and pseudo-reflection pulses are compared with the direct pulse shape. The direct pulse is the step in the radial displacement potential. The method used is a variation of Cagniard's work together with suitable approximations. The results are contained in Cagniard's original work. The present treatment is simpler than Cagniard's and better adapted to direct numerical calculations for many important cases. A numerical example is given.

Introduction. In a previous paper [Dix, 1961] work of Cagniard [1939] was modified so that numerical results could be obtained simply for reflected pulses and transmitted pulses well on the critical cone.

We assume two semi-infinite, elastic, homogeneous, isotropic media joined (welded) everywhere on an infinite plane. The point source is symmetrical step in the potential for the radial displacement.

This paper extends the class of easily soluble analytical problems to include head pulses, ordinary reflections beyond the critical angle, pseudo-reflections. The pseudo-reflection is a regular contribution centered at the arrival time for the ordinary reflection. The approximate shape of the pseudo-reflection can be constructed by passing the direct pulse through a differentiating network with poor high-frequency characteristics. It is ideally a logarithmic spike instead of a δ -function spike. Ordinary reflection beyond the critical cone has the same shape as the reflection inside the critical cone except for a nearly negligible modification. The head pulse shape is approximately that of the integral of the direct wave. We follow the procedures and notations, with additions, used in the previous paper [Dix, 1961].

General considerations. We assume $r > r_c$ and large distance; i.e., the receiver is outside of the critical cone. Also we study only the vertical meridional part of the displacement, that is, the case $j = 1$ [Dix, 1961, equation 13].

Thus we have to consider

$$C_1(r, z, t) = I(t - t_{F1}) \frac{2S}{\pi h} \cdot \text{Im} \int_{E(1, \tau)} \frac{f(w)v \, dw}{[v^2 \eta^2 + \{\tau - (1 + \zeta)\alpha\}^2]^{1/2}} \quad (1)$$

Here $E(1, \tau)$ is the map of $K(1, t)$ rather than of $k(1, t)$. $E(1, \tau)$ begins at the point in the v plane where $\tau = i\eta v + (1 + \zeta)\alpha$. This point will be on the negative imaginary axis for $\tau \leq R_0$ and in the fourth quadrant for $\tau > R_0$. When $\tau < R_0$, $E(1, \tau)$ starts at $-i[\eta\tau - (1 + \zeta)(R_0^2 - \tau^2)^{1/2}]/R_0^2$ and goes up the axis to $-iJ$. When $\tau > R_0$ the same formulas for the end points apply, and the beginning point is $(1 + \zeta)(\tau^2 - R_0^2)^{1/2}/R_0^2 - i\eta\tau/R_0^2$. From this $E(1, \tau)$ goes to $-i\eta/R_0$ and then up the axis to $-iJ$.

Transfer to the α plane. Then (1) becomes

$$C_1(r, z, t) = I(t - t_{F1}) \frac{2S}{\pi R} \cdot \text{Im} \int_{H(1, \tau)} \frac{(G_5 + \alpha' G_6)\alpha \, d\alpha}{(\alpha^2 - 2\alpha_1\alpha + \alpha_1^2 + \alpha_2^2)^{1/2}} \quad (2)$$

where

$$f = G_5 + \alpha' G_6,$$

$$G_5 = (G_1 G_3 - G_2 G_4 \alpha'^2)/(G_3^2 - G_4^2 \alpha'^2),$$

$$G_6 = (G_2 G_3 - G_1 G_4)/(G_3^2 - G_4^2 \alpha'^2),$$

$$G_1 = -(w - \gamma B)^2 \alpha \beta - \gamma B^2 \alpha \beta' - w[w + B(1 - \gamma)]^2,$$

$$G_2 = w \alpha \beta \beta' + (w + B)^2 \beta' + \gamma B^2 \beta,$$

¹Contribution 998, Division of Geological Sciences, California Institute of Technology.

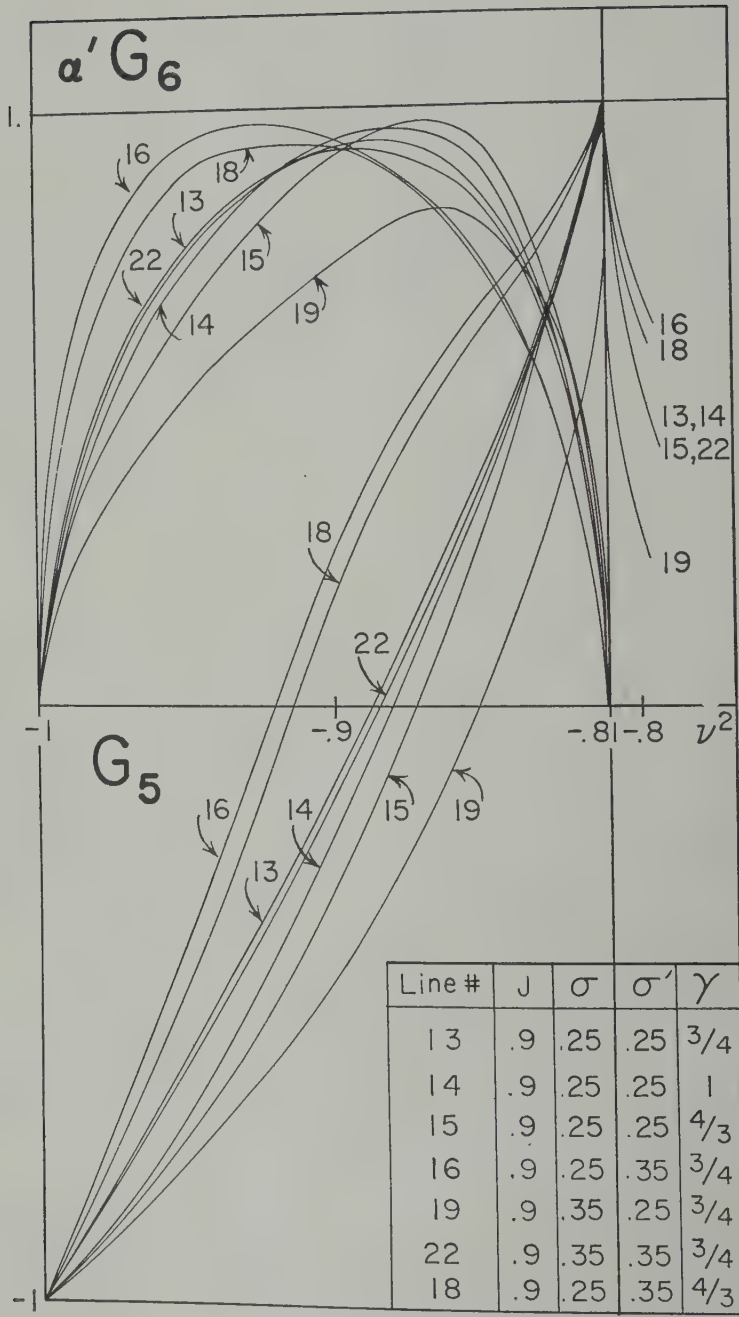


Fig. 1. Real G_5 and imaginary $-\alpha'G_6$ parts of the reflection function for longitudinal waves beyond the critical cone for the case where the velocity ratio is 0.9.

$$\begin{aligned}
 &= G_1 + 2w[w + B(1 - \gamma)]^2, \\
 &= 2w\alpha\beta\beta' - G_2, \\
 &= \alpha_1 - i\alpha_2, \quad \alpha_1 = (1 + \zeta)\tau/R_0^2, \\
 &= \eta(\tau^2 - R_0^2)^{1/2}/R_0^2.
 \end{aligned}$$

the head pulse. Here $\tau < R_0$, so α_0 is positive. In (2), G_6 is real, so it contributes nothing. α is also real, but α' is negative imaginary. In (2) becomes

$$\begin{aligned}
 r, z, t) &= -1(t - t_{F1}) \frac{2S}{\pi R} \\
 &\int_{\alpha_0}^{(1-J^2)^{1/2}} \frac{\alpha[(1-J^2)^{1/2} + \alpha]^{1/2} G_6}{(\alpha - \alpha_1 - i\alpha_2)^{1/2}} \\
 &\frac{(1-J^2)^{1/2} - \alpha]^{1/2}}{(\alpha - \alpha_0)^{1/2}} d\alpha \quad (3)
 \end{aligned}$$

first term in the integrand is slowly varying over the interval $\alpha_0 < \alpha < (1 - J^2)^{1/2}$ and the second term is singular at each end, being infinite at α_0 and zero at $(1 - J^2)^{1/2}$. Equation 3 is exact.

It seems hopeless to try to integrate (3) analytically. However, the first term in the integrand can be computed numerically for various cases, and it is seen that its variation is small (if we are not too close to the reflection time, SR !). We can use the mean value theorem and

$$\begin{aligned}
 &\frac{[(1 - J^2)^{1/2} - \alpha]^{1/2}}{(\alpha - \alpha_1 + i\alpha_2)^{1/2}} d\alpha \\
 &= \frac{\pi}{2} [(1 - J^2)^{1/2} - \alpha_0] \quad (4)
 \end{aligned}$$

$$\begin{aligned}
 r, z, t) &= -1(t - t_{F1}) \frac{S[(1 - J^2)^{1/2} - \alpha_0]}{R} \\
 &\text{mean} \left[\frac{\alpha[(1 - J^2)^{1/2} + \alpha]^{1/2} G_6}{(\alpha - \alpha_1 - i\alpha_2)^{1/2}} \right] \quad (5)
 \end{aligned}$$

Notice that at $t = t_{F1}$, $\alpha_0 = (1 - J^2)^{1/2}$, so $r, z, t_{F1}) = 0$ in the limit as $t \rightarrow t_{F1}$ from values of t such that $t > t_{F1}$. Thus the graph of $r, z, t)$ starts at zero at $t = t_{F1}$. Since G_6 is positive, $C_1(r, z, t)$ has a positive slope at t_{F1} . Consequently L_{sL} is a negative step function to a first approximation.

What is the variation of the step in L_{sL} with r ? For this we have to calculate

$$\begin{aligned}
 &\lim_{t \rightarrow t_{F1}} \frac{1}{Rh} \left\{ \frac{\partial \alpha_0}{\partial \tau} \right. \\
 &\quad \cdot \text{mean} \left[\frac{\alpha[(1 - J^2)^{1/2} + \alpha]^{1/2} G_6}{(\alpha - \alpha_1 - i\alpha_2)^{1/2}} \right] \\
 &\quad \left. - [(1 - J^2)^{1/2} - \alpha_0] \frac{\partial}{\partial \tau} \right. \\
 &\quad \cdot \text{mean} \left[\frac{\alpha[(1 - J^2)^{1/2} + \alpha]^{1/2} G_6}{(\alpha - \alpha_1 - i\alpha_2)^{1/2}} \right] \left. \right\} \quad (6)
 \end{aligned}$$

As $\lim [(1 - J^2)^{1/2} - \alpha_0] = 0$ and as its factor is bounded, the second term in the braces of (6) is zero. Also

$$\begin{aligned}
 \frac{\partial \alpha_0}{\partial \tau} &= \frac{-J}{(\eta(1 - J^2)^{1/2} - (1 + \zeta)J)} \\
 &= \frac{-J}{(\eta - \eta_c)(1 - J^2)^{1/2}} \quad (7)
 \end{aligned}$$

where $\eta_c = r_c/h = (1 + \zeta)J/(1 - J^2)^{1/2}$. Also

$$\begin{aligned}
 &\lim \text{mean} \left[\frac{\alpha[(1 - J^2)^{1/2} + \alpha]^{1/2} G_6}{(\alpha - \alpha_1 - i\alpha_2)^{1/2}} \right] \\
 &= \frac{(1 - J^2)^{1/2} R_0 (G_6)_{\alpha_0}}{\eta^{1/2} (\eta - \eta_c)^{1/2}} \quad (8)
 \end{aligned}$$

So the step is

$$-\{(G_6)_{\alpha_0} - (1 - J^2)^{1/2}\} J/r^{1/2} (r - r_c)^{3/2} \quad (9)$$

Near $r = r_c$ the step varies as $(r - r_c)^{-3/2}$. For very large r the variation is as r^{-2} .

The ordinary reflection beyond the critical distance. This phase is contributed by G_6 in (2) for the case $\tau > R_0$. It must be remembered that $\alpha_0 = \alpha_1 - i\alpha_2$ is complex since both α_1 and α_2 are real and positive. G_6 may be expanded about $\alpha = (1 + \zeta)/R_0$ on the real axis in the α plane. Then the problem is handled in just the same way as for the reflection for $r \ll r_c$ [Dix, 1961].

To facilitate numerical work the graphs of Figures 1 and 2 are included.

The pseudo reflection. This phase is contributed by $\alpha' G_6$ in (2) for τ in the neighborhood of R_0 . We consider τ close but not equal to R_0 .

Case $\tau < R_0$. $H(1, \tau)$ is a path similar to that used for the head pulse. However, it is not now legitimate to separate the two factors in the radical denominator of (2). Equation 2 may

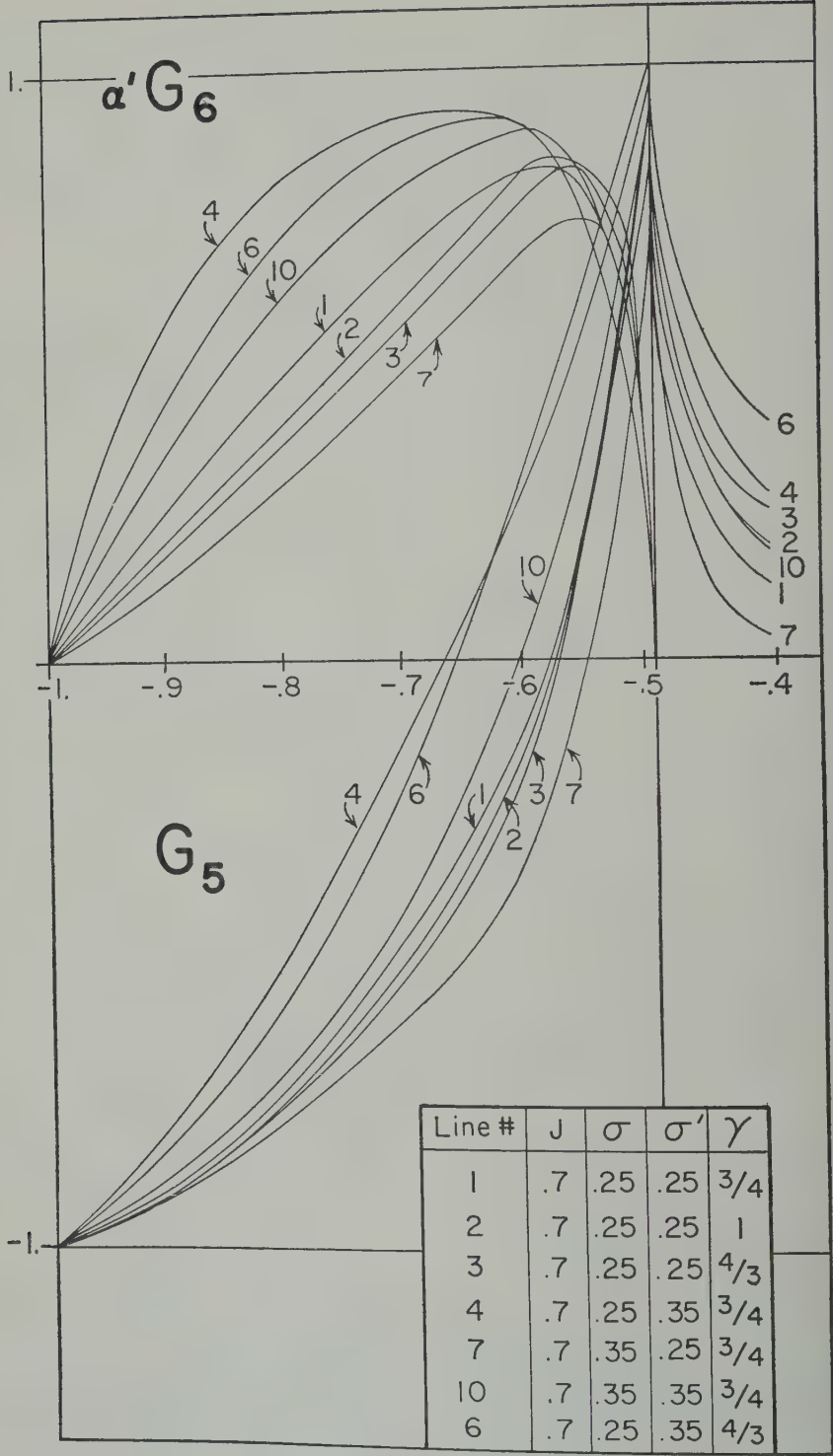
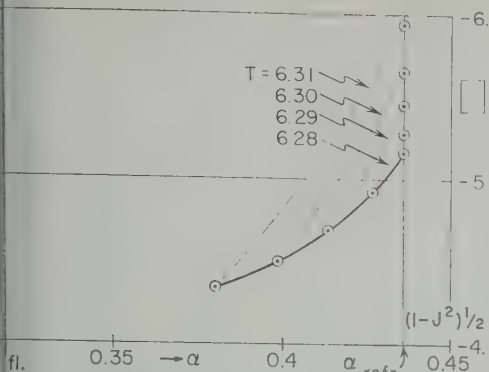


Fig. 2. Same as Figure 1 except for a velocity ratio of 0.7.



3. Graphs of the bracketed expression in equation 5.

Written

τ, z, t

$$-1(t - t_{F1}) \frac{2S}{\pi R} \text{mean}(|\alpha'| G_0)$$

$$\int_{\alpha_0}^{(1-J^2)^{1/2}} \frac{\alpha d\alpha}{(\alpha^2 - 2\alpha_1\alpha + \alpha_1^2 + \alpha_2^2)^{1/2}} \quad (10)$$

integral is

$$\begin{aligned} & -J^2 - 2\alpha_1(1 - J^2)^{1/2} + \alpha_1^2 + \alpha_2^2)^{1/2} \\ & \alpha_1 \ln \{ [1 - J^2 - 2\alpha_1(1 - J^2)^{1/2} \\ & \alpha_1^2 + \alpha_2^2]^{1/2} + (1 - J^2)^{1/2} - \alpha_1 \} \\ & \alpha_1 \ln(-i\alpha_2) \end{aligned} \quad (11)$$

first two terms remain bounded but $-\alpha_1$ is singular at $\tau = R_0$.

use $\tau > R_0$. For this $H(1, \tau)$ starts at $(1 + \zeta)\tau/R_0^2 - i\eta(\tau^2 - R_0^2)^{1/2}/R_0^2$ in the fourth quadrant of the α plane, goes up to $(1 + \zeta)/R_0$, and then goes along the real axis to $(1 - J^2)^{1/2}$.

consider the path from α_r to $(1 - J^2)^{1/2}$. G_5 contributes nothing. The integral we use is (10) with limit α_0 replaced by α_r . For $\tau > R_0$ the integrand is finite. The integration is carried out as before, and the first two terms are of the same form as the first two of (11). The remaining terms corresponding to the lower limit of integration are

$$\begin{aligned} & -2\alpha_1\alpha_r + \alpha_1^2 + \alpha_2^2)^{1/2} \\ & \alpha_1 \ln [(\alpha_r^2 - 2\alpha_1\alpha_r + \alpha_1^2 + \alpha_2^2)^{1/2} \\ & - \alpha_1] \end{aligned} \quad (12)$$

It can be shown that the logarithmic terms in (11) and (12) are locally symmetrical about $\tau = R_0$. They have the same sign. If we formed the time derivative to get L_{sL} we would have to use the 'Cauchy principle part' of the convolution integral.

The path from α_0 to α_r contributes very little because $\alpha'G_5$ is nearly purely imaginary and so is $d\alpha$.

Modification of the ordinary reflection. The first term in (12) is a curious one. When differentiated with respect to τ it presents a $(\tau - R_0)^{-1/2}$ singularity. The way to handle this is to convolute first and then differentiate. In fact this procedure should be used generally for all parts of L_{sL} . Then we shall have (for reflection terms)

$$\begin{aligned} \frac{\partial}{\partial \tau} \int_{R_0}^{\tau} C_1(\tau') F'(\tau - \tau') d\tau' &= C_1(\tau) F'(0) \\ &+ \int_{R_0}^{\tau} C_1(\tau') F''(\tau - \tau') d\tau' \end{aligned} \quad (13)$$

For most physically possible inputs, $F'(0) = 0$. Also F'' is just as easy to integrate as F' .

This curious term contributes a true reflection part which is to be superposed on the ordinary reflection from G_5 . It has no forerunner.

Numerical example. This is a rough approximation to the problem of the base of the crust. The velocity corresponding to the crustal velocity is taken to be 7.3 km/sec and the subcrustal velocity is 8.1 km/sec. We suppose that Poisson's ratio is 1/4 in each medium and that the density ratio (ρ'/ρ) is 4/3. The source is taken to be 40 km from the interface, and the receiver position is also 40 km from the interface on the same side as the source. The critical horizontal distance for this case is 165 km, beyond which the head wave begins to appear. To separate the head wave clearly from the reflected waves, we consider a horizontal distance of 240 km. Since the time differentiation is most easily carried out in the convolution, we compute only C_1 .

Head pulse. We apply the result of (5) to this case, making the calculations partly with the help of a digital computer for G_5 and with the help of a slide rule for all of the other computations. The values of the bracketed expression in (5) are shown in Figure 3, for four times after the time of the beginning of the head pulse. It is seen from the figure that the variation of

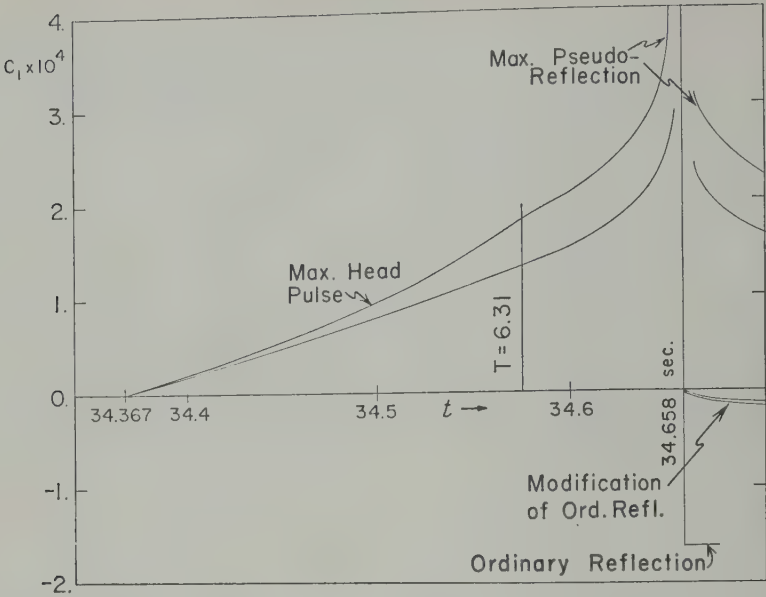


Fig. 4. Plot of the potential C_1 corresponding to the numerical example.

the bracketed expression is small and that the variation of the mean value is even smaller. We can, in fact, use the maximum value and the minimum value to get outside limits of the range of variation of C_1 . This we have shown in Figure 4. C_1 for the head pulse goes only from $t_{P1} = 34.367$ sec to $t = 34.658$ sec because we have decided to pass over to a consideration of the pseudo-reflection at this point, an arbitrary decision.

The only part of this calculation that is particularly difficult is the calculation of G_6 , and this is tabulated for several values of the elastic parameters in Table 1. From the table a fairly large class of numerical examples may be computed using a slide rule.

Ordinary reflection. This contribution gives for C_1 [Dix, 1961, equation 18], for the first approximation, a step (getting G_6 from Fig. 1),

$$(C_1)_{\text{REFL}} = 1(t - 34.658)$$

$$\begin{aligned} & \frac{1}{253 \times 7.3} [-0.95 \times 0.316] \\ &= -1(t - 34.658)(-1.62 \times 10^{-4}) \end{aligned}$$

This is plotted in Figure 4. It must be observed that the approximation, though exact for the size of the step, gets poor very rapidly as t increases, because we are so very close to the

branch singularity of G_6 at $w = -1$. The present approximation does give the δ function for i exactly, and the great distance usually makes the other terms negligible. The exception would correspond to the case in which the input spectrum contained appreciable low frequencies corresponding to waves of length comparable with r . By separating the α singularity from $f(w)$, as say $G_6 + \alpha G_{10}$, this case also can be handled.

The pseudo-reflection—Case $\tau < R_0$. We use (10) and (11). Reference to Figure 1 shows that $-|\alpha'| G_6$ varies from 0 to 0.98, so we cannot use (10) directly with extreme values of $|\alpha'| G_6$. We can, however, set limits by returning to the head-wave calculations. Note that (see Fig. 4) at $\tau = 6.31$ we have $\alpha_0 = 0.38$ and $C_{1\text{min}} = 1.32 \times 10^{-4}$ and $C_{1\text{max}} = 1.80 \times 10^{-4}$. In (10) we let mean $(|\alpha'| G_6) = X$ be undetermined. We then calculate C_1 using (10) and (11). This gives, at $\tau = 6.31$,

$$\begin{aligned} C_1 = & -3.45 \times 10^{-4} X \{0.1018 - 0.473 \\ & + 0.861\} = -1.69 \times 10^{-4} \end{aligned}$$

So $-X_{\text{max}} = 1.067$ and $-X_{\text{min}} = 0.778$. We use these factors for (11). The results are plotted in Figure 4 between $\tau = 6.31$ and $t = 34.658$ sec.

Case $\tau > R_0$. We assume the same determination of the factor X as above. This is necessary

TABLE 1

σ^1	γ	$G_{\theta, J=0.9}$	$G_{\theta, J=0.7}$
0.25	4/3	-6.29	-4.59
0.25	1	-5.77	-4.42
0.25	3/4	-4.75	-4.48
0.35	4/3	-3.72	-3.44
0.35	1	-3.89	-3.64
0.35	3/4	-3.90	-3.70
0.25	4/3	-8.05	-6.22
0.25	1	-7.21	-5.91
0.25	3/4	-6.14	-5.36
0.35	4/3	-5.18	-7.06
0.35	1	-5.69	-4.37
0.35	3/4	-5.32	-4.46

the symmetry about $\tau = R_0$. So we continue calculation of the first two terms of (11) and add to this the two terms of (12). The results are plotted in Figure 4.

Modification of the ordinary reflection. This modification comes from the bounded term in (11). This small contribution has been plotted in Figure 4.

Summary. In the preceding pages simple approximate methods have been presented for calculation of the main characteristics of the head pulse, the ordinary reflection beyond the critical angle, and the pseudo-reflection.

The 'ordinary reflection' is defined and occurs continuously from normal incidence right through the critical angle and on out to infinite distance. There is always a place beyond the critical angle where the pulse becomes zero and then at greater distances continues with a changed shape. The ordinary reflection has its maximum amplitude at the critical distance. This maximum is sharply peaked, so any lack of the ideal conditions (plane interface, homogeneous media) will seriously complicate this maximum. A comparison of Figure 1 suggests that the most important contrast in parameters may be that between Poisson's ratios for the two media. The ordinary reflection beyond the critical angle is usually largely a 'far-field' phenomenon and as such suffers very little change of shape. Except for very long wave effects, if absent, the calculation method given is as well as simple.

The shape of the head pulse is about the same as that of the near-field contribution of the incident pulse. Hence it is likely to be missed

unless the source contains plenty of long-wave energy and the receiver is more sensitive to the long waves than to the short ones. This gives the way to separate the head wave from the direct wave and from the ordinary reflection by filtering at the receiver.

The pseudo-reflection is centered at the ordinary reflection time. It must be understood in terms of processes within both media. Its character may be contrasted roughly with the ordinary reflection and with the head wave as follows: Let $F(t)$ be the input time variation of the source. Then the head wave corresponds to $\int_0^t F(t') dt'$, the ordinary reflection to $F(t)$, and the pseudo-reflection very roughly to $-dF/dt$.

The numerical example can be used to provide a definite case of each item discussed above. For example, the ordinary reflection changes sign at about 210 km horizontal distance from the source.

There is very little novelty in the above presentation so far as theory is concerned. It only arranges the theory for simple approximate calculation.

There is a fairly extensive collection of papers related to this subject [Jeffreys, 1926; Wolf, 1936; Heelan, 1953; Donato, 1960; Ewing, Jardetzky, and Press, 1957]. For a good brief survey of the Russian literature see Bessonova and others [1960, pp. 42-48] and references contained there. This is a partial list.

REFERENCES

- Bessonova, E. N., and others, *Investigation of the Mechanism of Earthquakes*, Trans. (Trudy) Geophys. Inst., Acad. Sci. USSR, no. 40 (1966), English translation publ. by Am. Geophys. Union as *Soviet Research in Geophysics*, vol. 4, 1960.
- Cagniard, L., *Reflexion et Refraction des Ondes Seismiques Progressives*, Gauthier-Villars, Paris, 1939.
- Dix, C. H., The reflected seismic pulse, *J. Geophys. Research*, 66, 227-233, 1961.
- Donato, R. J., *Geophys. J.*, 3, 270-271, 1960.
- Ewing, M., W. S. Jardetzky, and F. Press, *Elastic Waves in Layered Media*, McGraw-Hill Book Co., New York, 1957.
- Heelan, P. A., *Geophysics*, 18, 871-893, 1953.
- Jeffreys, H., *Proc. Cambridge Phil. Soc.*, 23, 472-481, 1926.
- Wolf, A., *Geophysics*, 1, 319-326, 1936.

(Manuscript received March 16, 1961; revised July 1, 1961.)

Elastic Wave Propagation in Layered Anisotropic Media¹

DON L. ANDERSON

*Seismological Laboratory, California Institute of Technology
Pasadena, California*

Abstract. This is an analysis of the dispersive properties of transversely isotropic media. This kind of anisotropy is exhibited by hexagonal crystals, sediments, planar igneous bodies, ice sheets, and rolled metal sheets where the unique axis is perpendicular to the direction of surface wave propagation and the other axes are distributed randomly in the plane of the layers. Period equations are derived for waves of Rayleigh, Stoneley, and Love types, and comparisons are made, in certain cases, with ray theoretical and plane stress solutions. Anisotropy can have a pronounced effect on both the range of existence and the shape of the dispersion curves and can lead to an apparent discrepancy between Love and Rayleigh wave data. Attention is focused in this initial paper on a single solid layer *in vacuo* (i.e. a free plate) and a solid layer in contact with a fluid halfspace. The single layer solutions are generalized to n -layer media by the use of Haskell matrices.

Introduction. Elastic wave problems are usually formulated for convenience under the restrictive assumptions of homogeneity, perfect continuity, plane parallel boundaries, and isotropy. Although these assumptions are often approximately satisfied in practice, certain ambiguities exist between theory and observation that create the need for reformulating some of these problems under less restrictive and possibly more realistic assumptions. Herein we drop the assumption of isotropy.

Isotropy is exhibited in its purest form in single crystals but also occurs in collections of grains or minerals which have crystallized or have been deposited with a preferred orientation, or have been subjected to nonuniform forces during formation. A layered medium by its very nature is anisotropic in the large, but the individual layers may also be anisotropic in a manner which cannot be handled by a further subdivision into thinner layers. Heterogeneous media with random grain orientation tend to be isotropic. In this paper we shall consider materials that possess an axis of symmetry in the sense that all rays at equal angles to this axis are equivalent. Such materials are called 'transversely isotropic,' and we shall consider wave propagation in an infinite or semi-infinite

medium with this symmetry has been discussed by Love [1944], Sató [1950], Musgrave [1959], and Stoneley [1949]. A convenient summary is given by Mason [1958] and Ewing, Jardetzky, and Press [1957]. Transverse isotropy results in the same set of elastic constants as that for hexagonal symmetry and hence is exhibited in all metals or minerals crystallizing in the hexagonal system. This symmetry is also expected to be displayed by sediments, planar igneous bodies, floating ice sheets, and rolled or extruded metal and plastic sheets. This latter material is often used in two-dimensional model experiments.

An isotropic solid is governed by two elastic constants and the characteristic equation has three roots, one corresponding to a compressional wave and a double root corresponding to a distortional wave. These velocities are independent of direction. A transversely isotropic solid is governed by five elastic constants, and the separation into two waves, one for which the curl of the displacement vanishes and one for which the divergence of the displacement vanishes, does not in general occur. Corresponding to any assigned wave normal there are three velocities of elastic wave propagation, and only in special cases do these degenerate to purely shear and purely compressional motion.

2. *Generalized Hooke's law and the equations of motion.* The matrix of the elastic constants for a medium with hexagonal or transverse

¹Contribution 1036, Division of Geological Sciences, California Institute of Technology, Pasadena, California.

isotropic symmetry is [Love, 1944]

$$\begin{array}{cccccc} c_{11} & c_{12} & c_{13} & 0 & 0 & 0 \\ c_{12} & c_{11} & c_{13} & 0 & 0 & 0 \\ c_{13} & c_{13} & c_{33} & 0 & 0 & 0 \\ 0 & 0 & 0 & c_{44} & 0 & 0 \\ 0 & 0 & 0 & 0 & c_{44} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{c_{11} - c_{12}}{2} \end{array}$$

For an isotropic body,

$$c_{11} = c_{33} = \lambda + 2\mu; \quad c_{13} = \lambda; \\ (c_{11} - c_{12})/2 = c_{44} = \mu$$

If we ignore body forces, the equations of small motion are three of the form,

$$\rho \frac{\partial^2 u}{\partial t^2} = \frac{\partial p_{xx}}{\partial x} + \frac{\partial p_{xy}}{\partial y} + \frac{\partial p_{xz}}{\partial z} \quad (1)$$

The stresses p_{ij} are derived from the strain energy function W by

$$p_{xx} = \frac{\partial W}{\partial e_{xx}}, \quad p_{xy} = \frac{\partial W}{\partial e_{xy}}, \quad \text{etc.} \quad (2)$$

where [Love, 1944]

$$2W = c_{11}(e_{xx}^2 + e_{yy}^2) + c_{33}e_{zz}^2 \\ + 2c_{13}(e_{xx} + e_{yy})e_{zz} + 2c_{12}e_{xx}e_{yy} \\ + c_{44}(e_{yz}^2 + e_{zx}^2) + \frac{1}{2}(c_{11} - c_{12})e_{xy}^2 \quad (3)$$

The stresses are, accordingly,

$$\begin{array}{l} p_{xx} = c_{11}e_{xx} + c_{12}e_{yy} + c_{13}e_{zz} \\ p_{yy} = c_{12}e_{xx} + c_{11}e_{yy} + c_{13}e_{zz} \\ p_{zz} = c_{13}(e_{xx} + e_{yy}) + c_{33}e_{zz} \\ p_{xy} = p_{yx} = \frac{1}{2}(c_{11} - c_{12})e_{xy} \\ p_{zy} = p_{yz} = c_{44}e_{yz} \\ p_{zx} = p_{xz} = c_{44}e_{xz} \end{array} \quad (2a)$$

From the symmetry of the above equations it is

obvious that z has been taken as the unique axis.

The equations of motion become

$$\begin{array}{l} \rho \frac{\partial^2 u}{\partial t^2} = \frac{\partial}{\partial x} \left(c_{11} \frac{\partial u}{\partial x} + c_{12} \frac{\partial v}{\partial y} + c_{13} \frac{\partial w}{\partial z} \right) \\ + \frac{\partial}{\partial y} \left(\frac{c_{11} - c_{12}}{2} \right) \frac{\partial u}{\partial y} + \frac{\partial}{\partial z} c_{44} \frac{\partial w}{\partial z} \\ \rho \frac{\partial^2 v}{\partial t^2} = \left(\frac{c_{11} - c_{12}}{2} \right) \frac{\partial^2 w}{\partial x \partial y} + \frac{\partial}{\partial y} \left\{ c_{12} \frac{\partial^2 u}{\partial x^2} \right. \\ \left. + c_{11} \frac{\partial^2 v}{\partial y^2} + c_{13} \frac{\partial^2 w}{\partial z^2} \right\} + c_{44} \frac{\partial^2 v}{\partial z^2} \\ \rho \frac{\partial^2 w}{\partial t^2} = c_{44} \frac{\partial^2 w}{\partial x^2} + c_{44} \frac{\partial^2 v}{\partial y \partial z} \\ + \frac{\partial}{\partial z} \left(c_{13} \frac{\partial^2 u}{\partial x^2} + c_{13} \frac{\partial^2 v}{\partial y^2} + c_{33} \frac{\partial^2 w}{\partial z^2} \right) \end{array}$$

where u, v, w are the displacements in the x, y, z directions. There is no advantage in introducing the standard potentials since the equations of motion are still unseparable.

3. *Plane waves in an infinite medium.* The theory of plane wave propagation in the interior of an infinite anisotropic body is well developed (see, for instance, Love [1944] or Mason [1958]). We begin with a brief review of this theory to establish the setting for the following section.

For plane waves propagated in a direction specified by direction cosines (l, m, n) we take

$$(u, v, w) = (U, V, W)e^{i\omega t - ik(lx + my + nz)}$$

Substitution into the equations of motion gives

$$\begin{bmatrix} \mathcal{F} - \rho c^2 & lm \left(\frac{c_{11} + c_{12}}{2} \right) & nl(c_{13} + c_{44}) \\ lm \left(\frac{c_{11} + c_{12}}{2} \right) & \mathcal{G} - \rho c^2 & mn(c_{13} + c_{44}) \\ nl(c_{13} + c_{44}) & mn(c_{13} + c_{44}) & \mathcal{H} - \rho c^2 \end{bmatrix} \begin{bmatrix} U \\ V \\ W \end{bmatrix} = 0$$

where

$$\begin{array}{l} \mathcal{F} = l^2 c_{11} + m^2 \left(\frac{c_{11} - c_{12}}{2} \right) + n^2 c_{44} \\ \mathcal{G} = l^2 \left(\frac{c_{11} - c_{12}}{2} \right) + m^2 c_{11} + n^2 c_{44} \\ \mathcal{H} = (l^2 + m^2) c_{44} + n^2 c_{33} \end{array}$$

By setting the determinant of the coefficients

al to zero, we obtain the velocity equation. special cases may be dealt with immediately:

b) For transmission along the unique axis, $l = 1, m = l = 0, c^2 = c_{33}/\rho$, and $c^2 = c_{44}/\rho$ solutions. The first corresponds to a vertically traveling purely compressional wave (PV) and the second is a double root corresponding to a vertically traveling shear wave with horizontal particle motion. The degeneracy is caused by SV and SH waves becoming indistinguish-

c) For transmission along the x or y direction or any other direction perpendicular to the unique axis, $n = 0$, the solutions are

$$c^2 = \frac{c_{11}}{\rho} \quad \text{compressional, } PH$$

$$c^2 = \frac{c_{44}}{\rho} \quad \text{shear, } SV$$

$$c^2 = \frac{c_{11} - c_{12}}{2\rho} \quad \text{shear, } SH$$

Therefore, measurement along these two directions will determine four of the five elastic constants. To determine the fifth we need a measurement at some intermediate angle. In a particular case we can set $l = n = 1/\sqrt{2}, m = 0$, and from the velocity equation obtain

$$c^2 = \left\{ [2\rho c^2 - \frac{1}{2}(c_{11} + c_{33} + 2c_{44})]^2 - \frac{1}{2}(c_{11} - c_{33})^2 \right\}^{1/2} - c_{44}$$

Expressing c_{13} in terms of the velocity of the fastest traveling wave at 45° to the z axis. Solving the velocity equation for arbitrary l, m, n , we can determine the directional dependence of the velocities.

In the following we shall use the designations

$$\alpha_1^2 = \alpha_2^2 \quad (PV); \quad c_{11}/\rho = \alpha_1^2 \quad (PH)$$

$$c_{44}/\rho = \beta_1^2 \quad (SH_V, SV_V, SV_H); \quad (7)$$

$$(c_{11} - c_{12})/2\rho = \beta_2^2 \quad (SH_H)$$

In an isotropic body

$$\alpha_1^2 = \alpha_2^2 = (\lambda + 2\mu)/\rho$$

$$\beta_1^2 = \beta_2^2 = \mu/\rho$$

Surface waves in an anisotropic layer. Consider now a layer of thickness $2H$ with the above

symmetry overlying a fluid halfspace with constants ρ_2, λ_2 . Take z increasing downward from the center of the layer. This configuration will permit us to study the effect of anisotropy in a relatively simple system for which the isotropic theory is well developed and for which experimental data are available. Also, with this general case in hand we can easily study as special cases the effect of anisotropy on Rayleigh and Stoneley waves and on propagation in a free plate. Later we shall discuss the general n -layer anisotropic problem and point out how anisotropy will introduce apparent discrepancies between Love and Rayleigh wave data as well as giving erroneous results for Love or Rayleigh data used alone. Since we shall be interested in applying the results of our present restricted problem to a high-speed layer overlying a low-speed fluid halfspace (the floating ice sheet problem) we have the additional problem of leakage for all modes with phase velocities greater than the fluid velocity, but this is resolved by programming our resultant period equation in complex algebra, thus permitting the location of complex roots. This, however, introduces no additional difficulties into our present analysis.

Restricting ourselves to motion in two dimensions (x, z) we put $\partial/\partial y = 0, e_{yy} = 0, e_{yz} = 0$.

For surface waves we seek solutions of the type

$$(u, w) = [U_i(z), W_i(z)]e^{i(\omega t - kx)} \quad (8)$$

Substitution into the equations of motion yields

$$\begin{aligned} -\rho_1\omega^2 U(z) &= -c_{11}k^2 U(z) \\ &\quad - ik(c_{13} + c_{44})W'(z) + c_{44}U''(z) \end{aligned} \quad (9)$$

$$\begin{aligned} -\rho_1\omega^2 W(z) &= c_{33}W''(z) \\ &\quad - ik(c_{13} + c_{44})U'(z) - k^2 c_{44}W(z) \end{aligned}$$

where the primes denote $\partial/\partial z$.

If $U(z) = Ue^{\nu_i z}$ and $W(z) = iWe^{\nu_i z}$, then

$$-\rho_1\omega^2 U = -c_{11}U + k\nu_i G W + c_{44}\nu_i^2 U \quad (10)$$

$$-\rho_1\omega^2 W = c_{33}\nu_i^2 W - k\nu_i G U - k^2 c_{44}W$$

where $G = c_{13} + c_{44}$.

The condition that nonzero roots of equation 10 exist is

$$\begin{aligned} (c_{44}\nu_i^2 - c_{11}k^2 + \rho_1\omega^2)(c_{33}\nu_i^2 - k^2 c_{44} + \rho_1\omega^2) \\ + k^2 \nu_i^2 G^2 = 0 \end{aligned} \quad (11)$$

The explicit values of ν^2 are

$$\nu_i^2 = -\frac{M_1}{2c_{33}c_{44}} \pm \frac{M_3}{2c_{33}c_{44}} \quad (12)$$

where

$$M_3^2 = M_1^2 - 4M_2c_{33}c_{44}.$$

$$M_1 = c_{33}(\rho_1\omega^2 - c_{11}k^2)$$

$$+ c_{44}(\rho_1\omega^2 - c_{44}k^2) + k^2G^2.$$

$$M_2 = (\rho_1\omega^2 - c_{11}k^2)(\rho_1\omega^2 - c_{44}k^2).$$

For an isotropic body equation 12 becomes

$$\nu_1^2 = (k^2 - \omega^2/\alpha^2), \quad \nu_2^2 = (k^2 - \omega^2/\beta^2),$$

$$\alpha^2 = (\lambda + 2\mu)/\rho, \quad \beta^2 = \mu/\rho$$

so that ν_1 reduces to the form associated with a pure compressional wave and ν_2 reduces to that for a pure distortional wave.

For a given ν_i the displacement ratios in the

in the solid, and

$$U^* = U_0 e^{-\nu' z} \quad (1)$$

$$W^* = -(i\nu'/k) U_0 e^{-\nu' z}$$

in the liquid, where $\nu'^2 = (k^2 - \omega^2/\alpha'^2)$; $\alpha'^2 = \lambda_2/\rho_2$.

The boundary conditions are

$$p_{zz} = c_{33} \frac{\partial W}{\partial z} + c_{13} \frac{\partial U}{\partial x} = 0; \quad z = -H$$

$$p_{zx} = c_{44} \left(\frac{\partial U}{\partial z} + \frac{\partial W}{\partial x} \right) = 0; \quad z = -H$$

$$p_{xx} = 0; \quad p_{zz} = p_{zz}^*; \quad W = W^*; \quad z = H$$

Substituting equations 14 into equations gives

$$U_1[-\gamma_1\nu_1c_{33} + c_{13}k]sh\nu_1H + U_2[\gamma_1\nu_1c_{33} - c_{13}k]ch\nu_1H + U_3[-\nu_2\gamma_2c_{33} + c_{13}k]sh\nu_2H \\ + U_4[\nu_2\gamma_2c_{33} - c_{13}k]ch\nu_2H = 0$$

$$U_1[\nu_1 + k\gamma_1]ch\nu_1H + U_2[-\nu_1 - k\gamma_1]sh\nu_1H + U_3[\nu_2 + \gamma_2k]ch\nu_2H + U_4[-\nu_2 - \gamma_2k]sh\nu_2H =$$

$$U_1[\gamma_1\nu_1c_{33} - c_{13}k]sh\nu_1H + U_2[c_{33}\gamma_1\nu_1 - c_{13}k]ch\nu_1H + U_3[\nu_2c_{33}\gamma_2 - c_{13}k]sh\nu_2H$$

$$+ U_4[c_{33}\nu_2\gamma_2 - c_{13}k]ch\nu_2H + U_0 \frac{\lambda_2}{k} [k^2 - \nu'^2] e^{-\nu'H} = 0 \quad ($$

$$U_1[\gamma_1k]ch\nu_1H + U_2k\gamma_1sh\nu_1H + U_3k\gamma_2ch\nu_2H + U_4k\gamma_2sh\nu_2H + U_0\nu'e^{-\nu'H} = 0$$

$$U_1[\nu_1 + k\gamma_1]ch\nu_1H + U_2[\nu_1 + k\gamma_1]sh\nu_1H + U_3[\nu_2 + k\gamma_2]ch\nu_2H + U_4[\nu_2 + k\gamma_2]sh\nu_2H =$$

solid are

$$\left(\frac{U}{W} \right)_i = -\frac{k\nu_i G}{(c_{44}\nu_i^2 - c_{11}k^2 + \rho_1\omega^2)} \\ = \frac{(c_{33}\nu_i^2 - k^2c_{44} + \rho_1\omega^2)}{k\nu_i G} = \frac{1}{\gamma_i} \quad (13)$$

$$W_i = \gamma_i U_i$$

For isotropic media $\gamma_1 = \nu_1/k$, $\gamma_2 = k/\nu_2$. We therefore take as our solutions

$$U = U_1sh\nu_1z + U_2ch\nu_1z$$

$$+ U_3sh\nu_2z + U_4ch\nu_2z \quad (14a)$$

$$W = i\gamma_1 U_1ch\nu_1z + i\gamma_1 U_2sh\nu_1z$$

$$+ i\gamma_2 U_3ch\nu_2z + i\gamma_2 U_4sh\nu_2z$$

The condition that the determinant of coefficients vanish is the period equation

$$\Pi_1\Omega_2sh\nu_1(\Pi_1\Gamma_2sh\nu_2ch\nu_1 - \Pi_2\Gamma_1ch\nu_2sh\nu_1)$$

$$+ \Lambda_2\Pi_2sh\nu_2(\Pi_2\Gamma_1sh\nu_1ch\nu_2$$

$$- \Pi_1\Gamma_2sh\nu_2ch\nu_1)$$

$$+ \Pi_1\Omega_1ch\nu_1(\Pi_1\Gamma_2ch\nu_2sh\nu_1$$

$$- \Pi_2\Gamma_1sh\nu_2ch\nu_1)$$

$$+ \Lambda_1\Pi_2ch\nu_2(\Pi_2\Gamma_1sh\nu_2ch\nu_1$$

$$- \Pi_1\Gamma_2sh\nu_1ch\nu_2) = 0 \quad ($$

where

$$\Gamma_1 = [-\gamma_1\nu_1c_{33} + c_{13}k]; \quad \Pi_1 = [\nu_1 + k\gamma_1]$$

$$\Gamma_2 = [-\nu_2\gamma_2c_{33} + c_{13}k]; \quad \Pi_2 = [\nu_2 + k\gamma_2]$$

$$\nu' \Gamma_1 s h \nu_1 + \lambda_2 \gamma_1 (k^2 - \nu_1'^2) c h \nu_1$$

$$\nu' \Gamma_1 c h \nu_1 + \lambda_2 \gamma_1 (k^2 - \nu_1'^2) s h \nu_1$$

$$\nu' \Gamma_2 s h \nu_2 + \lambda_2 \gamma_2 (k^2 - \nu_1'^2) c h \nu_2$$

$$\nu' \Gamma_2 c h \nu_2 + \lambda_2 \gamma_2 (k^2 - \nu_1'^2) s h \nu_2$$

in isotropic layer this yields the period equation given by *Press and Ewing* [1951].

The asymptotic form of equation 17 for short wavelengths is

$$-\Pi_2 \Gamma_1 [\nu' (\Pi_1 \Gamma_2 - \Pi_2 \Gamma_1) + \nu_2 (k^2 - \nu_1'^2) (\Pi_1 \gamma_2 - \Pi_2 \gamma_1)] = 0 \quad (17a)$$

It will later be shown that the first factor is the Rayleigh equation for anisotropic media and the second factor is the Stoneley equation for the interface between an anisotropic solid and a

free plate. If ρ_2 is set equal to zero, the solutions of the period equation are

$$\left(\frac{\nu_1 H}{\nu_2 H} \right)^{\pm 1} \frac{\nu_1 + k \gamma_1 [\nu_2 \gamma_2 c_{33} - c_{13} k]}{\gamma_1 \nu_1 c_{33} - c_{13} k [\nu_2 + \gamma_2 k]} = \frac{\Gamma_2 \Pi_1}{\Gamma_1 \Pi_2} \quad (19)$$

The upper sign corresponding to antisymmetrical waves, and the lower sign corresponding to symmetrical waves in a free plate.

In the long wavelength limit this becomes

$$[\nu_1/\nu_2]^{\pm 1} = \Gamma_2 \Pi_1 / \Gamma_1 \Pi_2 \quad (20)$$

For symmetrical waves this gives

$$(\rho \omega^2 / k^2) - (c_{11} c_{33} - c_{13}^2) = 0 \quad (20a)$$

We can also derive the long wavelength limit for antisymmetrical waves from plane stress theory. As before the xy plane parallel to the xz plane of the plane and propagation in the yz plane, we have

$$\partial p_{xz} / \partial x = \rho \partial^2 u / \partial t^2 \quad (21)$$

Eliminating $\partial u / \partial z$ from the first and third of equations 3 yields

$$p_{xx} = \frac{c_{11} c_{33} - c_{13}^2}{c_{33}} \frac{\partial u}{\partial x} \quad (22)$$

Therefore

$$\rho \frac{\partial^2 u}{\partial t^2} = \frac{c_{11} c_{33} - c_{13}^2}{c_{33}} \frac{\partial^2 u}{\partial x^2} \quad (23)$$

and the plate velocity c_p is given by

$$c_p^2 = (c_{11} c_{33} - c_{13}^2) / \rho c_{33} \quad (24)$$

in agreement with (20a).

This derivation will serve as a check on the low-frequency limit of our ensuing calculations. Note that the plate velocity does not depend only on the horizontal compressional and SV velocities, as we may have suspected for the long wave limit, but involves the constants in the vertical and intermediate directions as well. This can lead to plate velocities which are outside the range possible for an isotropic solid, a fact which holds true also for the rest of the dispersion curve. This is related to the fact that directional Poisson's ratios in an anisotropic solid can exceed 0.5.

For waves short compared with the layer thickness, the hyperbolic tangents can be replaced by unity, giving

$$\Gamma_2 \Pi_1 - \Gamma_1 \Pi_2 = 0 \quad (25)$$

for both symmetrical and antisymmetrical waves.

This is, again, the Rayleigh equation for anisotropic media. A more detailed discussion of the free anisotropic plate is in preparation (I. Abubakar, personal communication).

6. Rayleigh and Stoneley waves. Consider an anisotropic halfspace with the above properties. Assume solutions that decrease exponentially with depth (z is positive downward and the solid lies in the top halfspace):

$$U = U_1 e^{\nu_1 z} + U_3 e^{\nu_3 z} \quad (26)$$

$$W = i \gamma_1 U_1 e^{\nu_1 z} + i \gamma_2 U_3 e^{\nu_3 z}$$

The ν_i and γ_i are the same as were derived previously for the general case. Applying the condition of a traction free surface we obtain the period equation

$$R = -\Gamma_1 \Pi_2 + \Gamma_2 \Pi_1 = 0 \quad (27)$$

This is the Rayleigh equation for transversely isotropic media first derived by *Stoneley* [1949]. It reduces in the case of isotropy to

$$R = [\nu_2^2 + k^2]^2 - 4 \nu_1 \nu_2 k^2 = 0 \quad (28)$$

For Stoneley waves we assume motion dying off into both the fluid and the solid. Assume equation 26 for motion in the solid, and, for the fluid,

$$U^* = U_0 e^{-\nu^* z} \quad (29)$$

$$W^* = -(i \nu' / k) U_0 e^{-\nu^* z}$$

Applying the conditions of a stress free interface and continuity of vertical displacement, we obtain the Stoneley period equation,

$$\frac{\lambda_2(k^2 - \nu'^2)}{\nu'} = \frac{\Pi_1 \Gamma_2 - \Pi_2 \Gamma_1}{\gamma_1 \Pi_2 - \gamma_2 \Pi_1} \quad (30)$$

For isotropy this goes into

$$R + \delta = 0 \quad (31)$$

where

$$\delta = \frac{\rho_2 \alpha^{*2} \nu_1}{\rho_1 \beta_1^2 \nu'} (\nu_2^2 - k^2)(\nu'^2 - k^2) \quad (32)$$

This is the form of the Stoneley equation given by *Press and Ewing* [1951]. When $\rho_2 = 0$, we obtain the Rayleigh equation. Thus, as we indicated earlier, the high-frequency limit of the fundamental symmetric and antisymmetric modes of a free plate is the Rayleigh velocity. A floating plate has two branches of the fundamental mode, one corresponding to the Rayleigh velocity and the other to the Stoneley velocity.

7. *Love waves.* In a layered transversely isotropic medium Love waves exist independently of Rayleigh waves; that is, there is no coupling between motions of Rayleigh and Love types for waves propagating in a plane perpendicular to the unique axis. This is the lowest symmetry for which this coupling does not, in general, exist.

Considering a free layer and taking the axes as before, we see that the equations of motion can be satisfied by putting $u = w = 0$, $\partial/\partial y = 0$. For waves of the Love type we take

$$v = V(z)e^{i(\omega t - kx)} \\ = [v_1 sh \zeta z + v_2 ch \zeta z]e^{i(\omega t - kx)} \quad (32)$$

and obtain the reduced wave equation

$$\partial^2 v / \partial t^2 = (N - \zeta^2 c^2) v k^2 / L \quad (33)$$

so that

$$\zeta^2 = (N - \rho c^2) k^2 / L \quad (34)$$

where $N = (c_{11} - c_{12})/2$, $L = c_{44}$. When $L = N = \mu = \rho \beta^2$, we have isotropy and

$$\zeta^2 = (k^2 - \omega^2 / \beta^2)$$

The boundary conditions are

$$\rho_{2v} = L \partial v / \partial z = 0 \quad \text{at } z = \pm H$$

These conditions lead to the period equation

$$\tanh 2kH[(N - \rho c^2)/L]^{1/2} = 0$$

If $\rho c^2 < N$, no roots of the above equation exist. Therefore we must have $\rho c^2 \geq N$, $c^2 \geq \beta_2^2$. The period equation can therefore be written

$$\tan 2kH(N/L)^{1/2}[(c^2/\beta_2^2) - 1]^{1/2} = 0$$

Equation 36 is satisfied if

$$2kH(N/L)^{1/2}[(c^2/\beta_2^2) - 1]^{1/2} \\ = n\pi \quad \text{where } n = 0, 1, 2 \dots$$

which is the Love period equation for an anisotropic plate.

This differs by a factor of $(N/L)^{1/2}$ from the isotropic Love wave period equation. Note that $c = \beta_2$ is always a solution. The above derivation also holds for an anisotropic layer in contact with a fluid layer on one or both sides.

As *Stoneley* [1949] has pointed out, it is the modulus c_{44} that resembles the isotropic rigidity for Rayleigh wave motion, whereas the corresponding modulus for Love wave motion is $(c_{11} - c_{12})/2$. This holds true for layered isotropic media in general and will make the velocity structure as determined for Rayleigh wave data inconsistent with Love wave information unless the anisotropy is taken into account.

For an anisotropic layer of thickness $2H$, the constants L_1 , N_1 , and β_2 over an anisotropic halfspace with constants L_2 , N_2 , and β_2 , we obtain for Love waves

$$\tan 2kH \left(\frac{c^2}{\beta_2^2} - 1 \right)^{1/2} \left[\frac{N_1}{L_1} \right]^{1/2} \\ = \left[\frac{L_2 N_2}{L_1 N_1} \right]^{1/2} \left[\frac{1 - (c^2/\beta_2^2)}{c^2/\beta_2^2 - 1} \right]^{1/2}$$

In the corresponding isotropic case $N = L$, and we recover the Love equation in familiar form. *Stoneley's* [1949] derivation of the isotropic Love equation is apparently in error.

It can be shown that the period equation expresses the condition of constructive interference between multireflected plane SH waves. This condition may be written, for the free

$$4H \cos \theta = n l_0$$

where θ is the angle the ray makes with the vertical and l_0 is the wavelength measured along the ray. The velocity of SH waves satisfies

$$\rho\beta^2(\theta) = l^2N + n^2L \quad (39)$$

With the substitutions $\beta(\theta)/c = \sin \theta$, $2\pi \sin \theta/l_0$, $\beta_1^2 = L/\rho$, and $\beta_2^2 = N/\rho$ we write for equations 38 and 39

$$n\pi = \frac{2kHc}{\beta(\theta)} \left[1 - \frac{\beta^2(\theta)}{c^2} \right]^{1/2} \quad (38a)$$

$$\beta^2(\theta) = \frac{L}{N} c^2 \left(\frac{c^2}{\beta_2^2} - 1 + \frac{L}{N} \right)^{-1} \quad (39a)$$

Substitution of (39a) into (38a) gives the equation

$$2kH(N/L)[(c^2/\beta_2^2) - 1]^{1/2} = n\pi \quad (40)$$

Numerical results. The general period equation (17) was programmed for the Cal Tech Computing Center's Burroughs 220 electronic computer, using a complex root-finding routine developed by *Phinney* [1961]. Modifications were built into the program, making it able to solve for the symmetrical and antisymmetrical modes in a free plate and to evaluate Rayleigh and Stoneley equations.

We have calculated the dispersion of Rayleigh waves in plates for three solids which have hexagonal symmetry: (a) beryl, (b) ice, and (c) a laminated solid. These results, of course, apply also to any anisotropic solid having the same relationship between the elastic constants as one of the above examples.

Beryl was chosen in order to extend Stoneley's results to a free plate. Ice was chosen as an example of a solid which exhibits a rather strong transverse isotropy and because lake ice commonly forms with the vertical c -axis (the unique axis) orientation. In the other axes are randomly oriented, but in beryl they are all equivalent the result is a plate having the properties of a single crystal. Sea ice and some lake ice form with a horizontal orientation. Here again we have a large plate with transverse isotropy, but the effective elastic constants are not the same as for a single crystal. A finely laminated solid

is transversely isotropic with the unique axis normal to the laminations; the effective elastic constants of a laminated solid must obey certain ordering rules which restricts the extent to which anisotropy may be approximated by layering. Table 1 gives the elastic constants used in computing these three cases.

To show more clearly the effect of anisotropy we have also calculated the dispersion for equivalent isotropic ice, that is $\alpha = \alpha_1$, $\beta = \beta_1$, and for some intermediate cases.

For convenience we introduce the following 'anisotropy factors':

$$\begin{aligned} \varphi &= c_{33}/c_{11} & \text{i.e. } \alpha_2^2 &= \varphi\alpha_1^2 \\ \xi &= (c_{11} - c_{12})/2c_{44} & \text{i.e. } \beta_1^2 &= \xi\beta_2^2 \\ \eta &= (c_{11} - 2c_{44})/c_{13} \end{aligned} \quad (41)$$

Isotropic media have $\phi = \xi = \eta = 1$.

Figure 1 shows the effect on the fundamental symmetric mode² (M_{11}) of increasing all of the anisotropy factors from below 1 to the values attained by anisotropic ice. For this kind of anisotropy, i.e. φ , ξ , and η increasing, the dispersion curve migrates uniformly toward higher phase velocity. Note the shift of the Airy phase. The low-frequency limit of M_{11} for an isotropic plate satisfies $c^2/\beta^2 = 4(1 - \beta^2/\alpha^2)$. Therefore $(c_p/\beta) \leq 2$. For an anisotropic plate with the constants of ice, (24) gives $(c_p/\beta_1) = 2.045$. This agrees with the numerical evaluation of the general period equation.

Stoneley [1949] has previously noted that the Rayleigh velocity for an anisotropic solid can be higher than that for a Poisson solid. We note, in addition, that it can be higher than is theoretically possible for an isotropic solid. This is true in particular for solids with the constants of beryl and ice. Therefore, both ends of M_{11} are greater than the theoretical maximum under the assumption of isotropy.

² We follow *Tolstoy and Usdin's* [1953] mode designation convention. See also *Ewing, Jardetzky, and Press* [1957, p. 283].

TABLE 1. Elastic and Anisotropic Parameters Adopted for Computation

c_{11}/c_{44}	c_{33}/c_{44}	c_{12}/c_{44}	c_{13}/c_{44}	φ	ξ	η	$\rho(\text{g/cm}^3)$
4.13	3.62	1.47	1.01	0.88	1.33	0.47	2.7
4.70	4.96	2.27	1.60	1.05	1.21	1.69	0.917
4.54	3.32	1.64	1.31	0.73	1.45	2.06	2.4

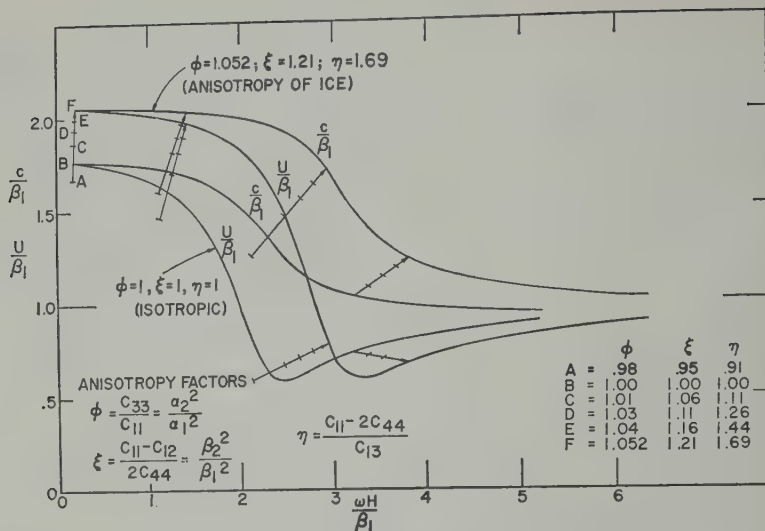


Fig. 1. Effect of anisotropy on M_{11} in a transversely isotropic plate.

To pursue this point further we have plotted in Figure 2 the variation in the phase velocities of the four lowest modes of an isotropic plate as Poisson's ratio changes from 0.25 to 0.5, the latter case indicating incompressibility. The data are from *Satō* [1951]. It is of interest to note that Poisson's ratio has a much greater effect on the symmetrical than on the anti-symmetrical modes. This is to be expected since we have effectively held β constant while changing α , and the M_1 modes are chiefly compressional in nature while the M_2 modes tend to couple with

an ideal free shear mode [*Tolstoy*, 1957]. Plotted for comparison are the four lowest modes for M_{11} and M_{11} for beryl. All the modes except M_{21} (flexural mode), which is not drawn, and the low-frequency end of M_{11} for beryl are outside the theoretical limit of isotropic plates. M_{21} for ice deviates slightly from isotropic values at both ends.

In Figure 3 are shown the complete results for ice, for M_{11} through M_{22} . Note the negative group velocity tail on M_{12} . This phenomenon has been discussed by *Tolstoy* [1957],

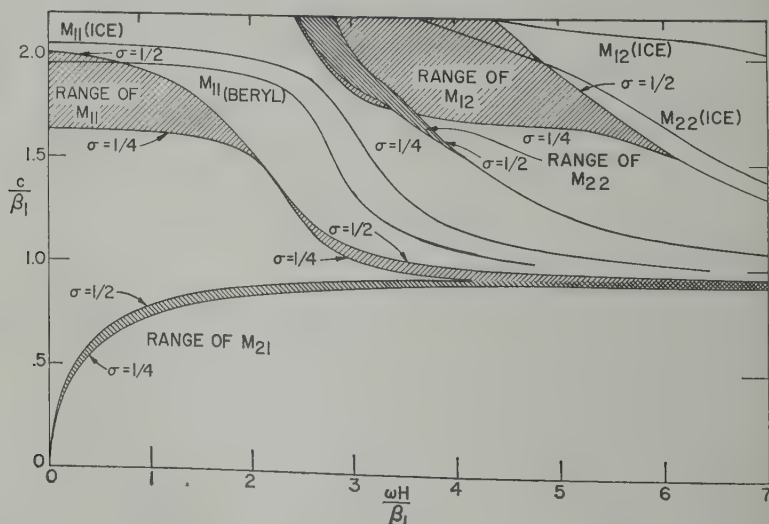


Fig. 2. Range of existence of M_{11} through M_{22} for $\frac{1}{4} < \sigma < \frac{1}{2}$ in isotropic plates and location of corresponding modes for anisotropic plates.

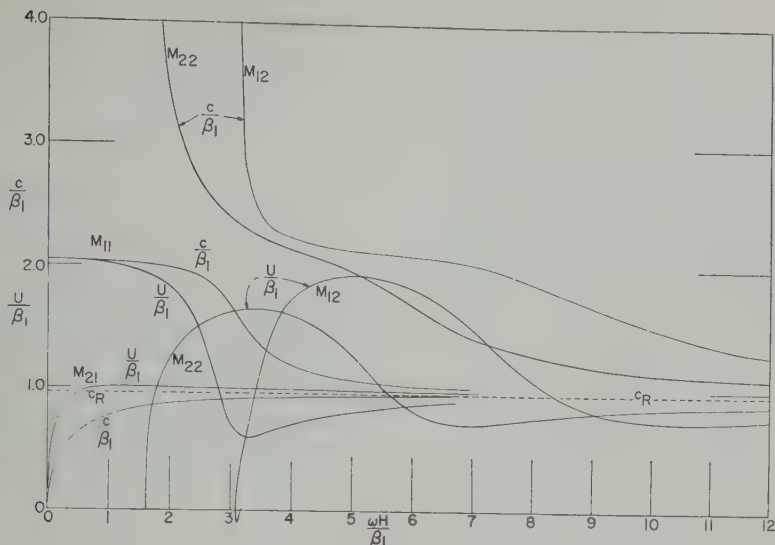


Fig. 3. Dispersion in a free anisotropic plate with the properties of ice.

routes it to a negative phase velocity rather than to an actual backward propagation of energy.

Figure 4 gives the results for a laminated plate consisting of alternating layers of sandstone and granite in the ratio 3 to 1. The effective elastic constants for waves long compared with the lamination thickness in this medium have been calculated by Postma [1955] and are listed in Table 1. The anisotropy factors are $\varphi = 0.733$, $\alpha^* = 1.450$, and $\eta = 2.0576$, making the anisotropy of this material quite different from that

of beryl or ice. Note the pulling down of the group velocity minimum and, in particular, the accentuated negative group velocity tail.

The Stoneley wave equation was evaluated numerically for $\rho_2/\rho_1 = 1.12$ (the water to ice density ratio) and several α^*/β_1 ratios. As in the isotropic case, the Stoneley velocity is always a fraction of the slower of α^* , β_1 , but for the constants tested it is a larger fraction than for the corresponding isotropic case. For example, when α^*/β_1 takes on the values 0.9, 1.0, 1.1, and 1.2, the Stoneley velocity

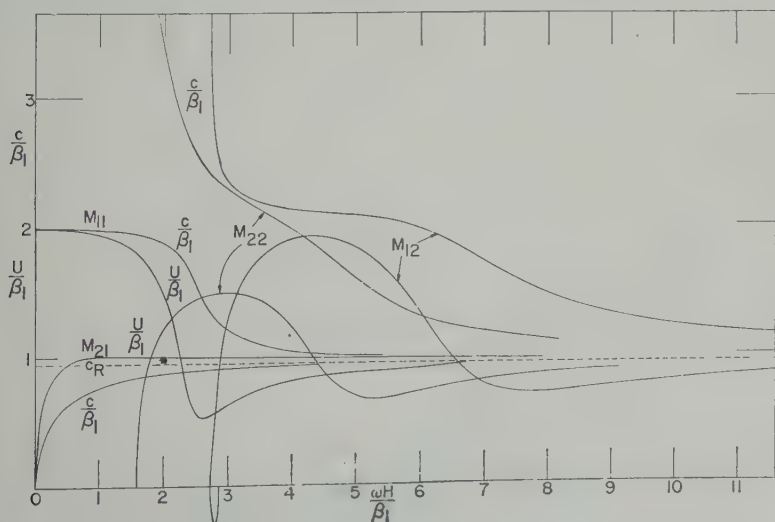


Fig. 4. Dispersion in a laminated plate.

c_{ST}/β_1 is 0.70, 0.74, 0.76, 0.78, and 0.79 for the anisotropic case and 0.67, 0.71, 0.72, 0.74, and 0.75 for the isotropic case.

9. *Matrix formulation of the general problem of a layered medium.* Following Haskell [1953] we now show how the solution of the wave propagation problem in a single anisotropic layer leads naturally to the n -layer problem. The normalized velocities and stresses in the m th layer can be written (understanding an $e^{i(\omega t - kx)}$ factor):

$$\begin{aligned} \dot{u}/c &= [iksh\nu_{1m}zU_{1m} + ikch\nu_{1m}zU_{2m} \\ &\quad + iksh\nu_{2m}U_{3m} + ikch\nu_{2m}zU_{4m}] \\ \dot{w}/c &= [-k\gamma_{1m}ch\nu_{1m}zU_{1m} - k\gamma_{1m}sh\nu_{1m}zU_{2m} \\ &\quad - k\gamma_{2m}ch\nu_{2m}zU_{3m} - k\gamma_{2m}sh\nu_{2m}zU_{4m}] \\ p_{zz} &= i(\gamma_{1m}\nu_{1m}C_m - F_mk)sh\nu_{1m}zU_{1m} \\ &\quad + i(C_m\gamma_{1m}\nu_{1m} - F_mk)ch\nu_{1m}zU_{2m} \\ &\quad + i(\gamma_{2m}\nu_{2m}C_m - F_mk)sh\nu_{2m}zU_{3m} \end{aligned}$$

$$+ i(\gamma_{2m}\nu_{2m}C_m - F_mk)ch\nu_{2m}zU_{4m}$$

$$\begin{aligned} p_{xz} &= L_m(\nu_{1m} + k\gamma_{1m})ch\nu_{1m}zU_{1m} \\ &\quad + L_m(\nu_{1m} + k\gamma_{1m})sh\nu_{1m}zU_{2m} \\ &\quad + L_m\nu_{2m} + \gamma_{2m}k)ch\nu_{2m}zU_{3m} \\ &\quad + L_m(\nu_{2m} + \gamma_{2m}k)sh\nu_{2m}zU_{4m} \end{aligned} \quad (10)$$

where C , L , and F are the elastic constants c_{44} and c_{13} . The boundary conditions to be used at each solid-solid interface are that these field quantities be continuous.

Taking $z = 0$ at the interface ($m - 1$), we can write the linear relationship between the motion stress vector and the displacement coefficients as

$$\begin{aligned} (\dot{u}_{m-1}/c, \dot{w}_{m-1}/c, p_{zzm-1}, p_{xzm-1}) \\ = E_m(U_m, V_m, W_m, Z_m) \end{aligned} \quad (11)$$

where $(U_1, U_2, U_3, U_4)_m = (U_m, V_m, W_m, Z_m)$ and where E_m is the matrix

$$E_m = \begin{bmatrix} 0 & ik & 0 & ik \\ -k\gamma_{1m} & 0 & -k\gamma_{2m} & 0 \\ 0 & i(C_m\gamma_{1m}\nu_{1m} - F_mk) & 0 & i(C_m\nu_{2m}\gamma_{2m} - F_mk) \\ L_m(\nu_{1m} + k\gamma_{1m}) & 0 & L_m(\nu_{2m} + k\gamma_{2m}) & 0 \end{bmatrix} \quad (12)$$

Setting $z = d_m$ we can write the relationship between the motion stress vector of the m th interface and the U_{im} :

$$(\dot{u}_m/c, \dot{w}_m/c, p_{zzm}, p_{xzm}) = D_m(U_m, V_m, W_m, Z_m) \quad (13)$$

where D_m is the matrix

$$D_m = \begin{bmatrix} iksh\nu_{1m}d_m & ikch\nu_{1m}d_m \\ -k\gamma_{1m}ch\nu_{1m}d_m & -k\gamma_{1m}sh\nu_{1m}d_m \\ \{i(\gamma_{1m}\nu_{1m}C_m - F_mk)sh\nu_{1m}d_m\} & \{i(C_m\gamma_{1m}\nu_{1m} - F_mk)ch\nu_{1m}d_m\} \\ L_m(\nu_{1m} + k\gamma_{1m})ch\nu_{1m}d_m & L_m(\nu_{1m} + k\gamma_{1m})sh\nu_{1m}d_m \\ iksh\nu_{2m}d_m & ikch\nu_{2m}d_m \\ -k\gamma_{2m}ch\nu_{2m}d_m & -k\gamma_{2m}sh\nu_{2m}d_m \\ \{i(\nu_{2m}\gamma_{2m}C_m - F_mk)sh\nu_{2m}d_m\} & \{i(C_m\nu_{2m}\gamma_{2m} - F_mk)ch\nu_{2m}d_m\} \\ L_m(\nu_{2m} + \gamma_{2m}k)ch\nu_{2m}d_m & L_m(\nu_{2m} + \gamma_{2m}k)sh\nu_{2m}d_m \end{bmatrix} \quad (14)$$

the coefficients U_i may be eliminated between equations 43 and 46, giving a linear relationship between the motion stress vector at the bottom of the top of the m th layer:

$$\begin{pmatrix} \dot{u}_m/c \\ \dot{w}_m/c \\ p_{zzm} \\ p_{zzm} \end{pmatrix} = D_m E_m^{-1} (\dot{u}_{m-1}/c, \dot{w}_{m-1}/c, p_{zzm-1}, p_{zzm-1}) \quad (47)$$

Applying (47) recursively we may carry our solution down through a stack of n layers. The development follows that of Haskell, and the properties of the matrices are the same as he discusses. By suitable alterations existing computer programs for isotropic layers can be modified to handle anisotropic layers.

Calculations for the n -layer problem will be presented in future papers.

Acknowledgments. I wish to thank Dr. Frank Press for his advice and encouragement throughout this work, and to acknowledge helpful discussions with R. Phinney, C. Archambeau, and D. Harkrider. This research was partially supported by contract 44-495 Ord. 1808 (Model Seismology) of the Army Research Office, Durham, and contract 63-638910 of the Air Force Technical Applications Center as part of the Advanced Research Projects Agency project Vela.

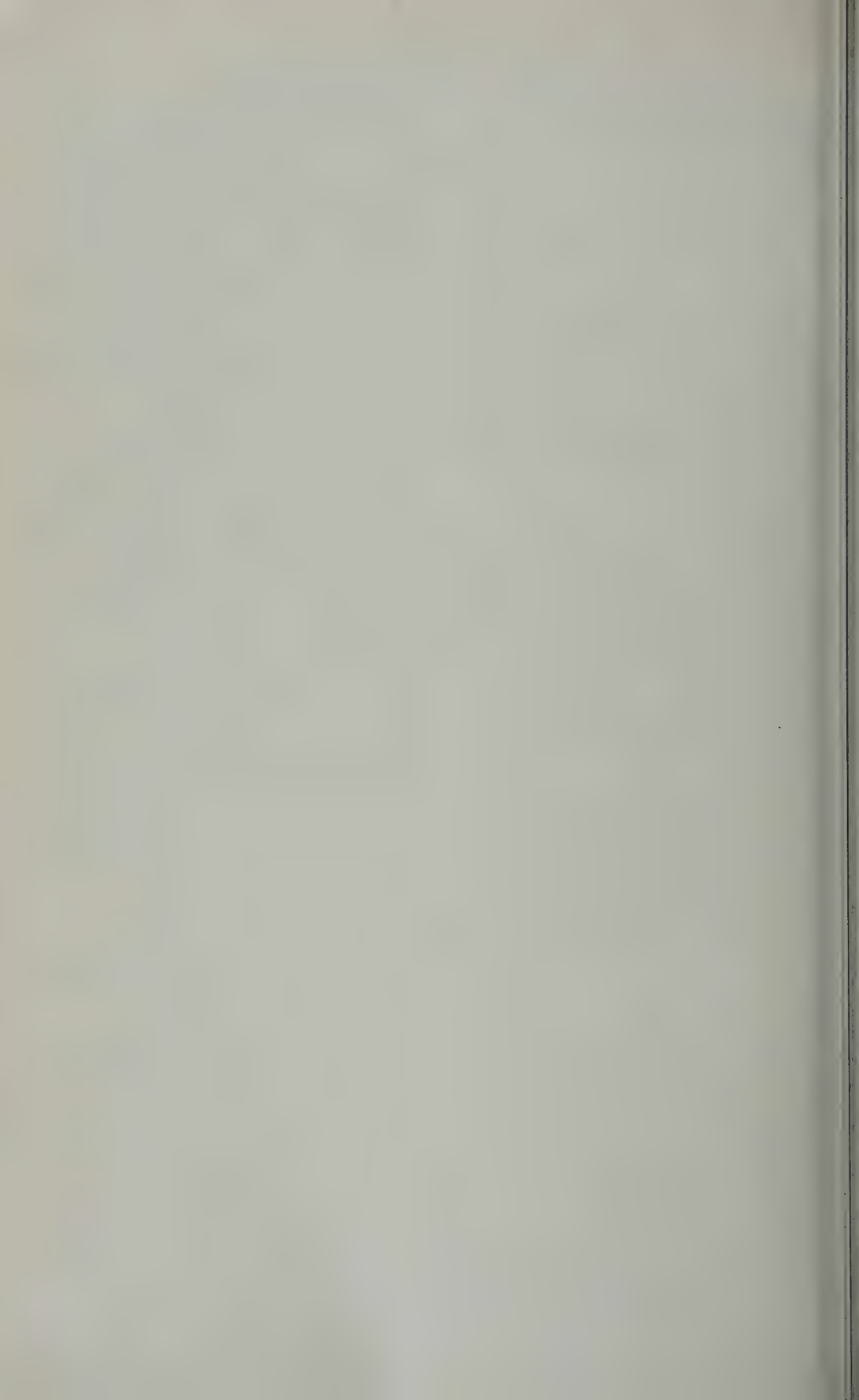
I gratefully acknowledge the support of a National Science Foundation predoctoral fellowship during the course of this work.

REFERENCES

Love, A. E. H., W. S. Jardetzky, and F. Press, *Elastic Waves in Layered Media*, McGraw-Hill Book Co., New York, 1957.

- Haskell, N. A., Dispersion of surface waves on multilayered media, *Bull. Seism. Soc. Am.*, 43, 17-34, 1953.
- Knopoff, L., On Rayleigh wave velocities, *Bull. Seism. Soc. Am.*, 42, 307-308, 1952.
- Love, A. E. H., *A Treatise on the Mathematical Theory of Elasticity*, 4th ed., reprinted by Dover Publications, New York, 1944.
- Mason, W. P., *Physical Acoustics and the Properties of Solids*, Bell Laboratory Series, D. Van Nostrand Co., Inc., Princeton, N. J., 1958.
- Musgrave, M. J. P., The propagation of elastic waves in crystals and other anisotropic media, *Repts. Progr. in Phys.*, 22, 74-96, 1959.
- Phinney, R. A., Leaking modes in the crustal waveguide, 1, The oceanic PL wave, *J. Geophys. Research*, 66, 1445-1469, 1961.
- Postma, G. W., Wave propagation in a stratified medium, *Geophysics*, 20, 780-806, 1955.
- Press, F., and M. Ewing, Propagation of elastic waves in a floating ice sheet, *Trans. Am. Geophys. Union*, 32, 673-678, 1951.
- Satô, Y., Rayleigh waves projected along the plane surface of a horizontally isotropic and vertically anisotropic elastic body, *Bull. Earthquake Research Inst. (Tokyo)* 28, 23-30, 1950.
- Satô, Y., Study on surface waves. II: Velocity of surface waves propagated upon elastic plates, *Bull. Earthquake Research Inst. (Tokyo)*, 29, 223-261, 1951.
- Stonely, R., The seismological implications of anisotropy in continental structure, *Monthly Notices Roy. Astron. Soc., Geophys. Suppl.* 5, 222-232, 1949.
- Tolstoy, I., and E. Usdin, Dispersive properties of stratified elastic and liquid media; a ray theory, *Geophysics*, 18, 844-870, 1953.
- Tolstoy, I., Wave propagation in elastic plates; low and high mode dispersion, *J. Acoust. Soc. Am.*, 29, 37-42, 1957.

(Manuscript received May 25, 1961.)



Phase and Group Velocities of Rayleigh Waves in a Spherical, Gravitating Earth¹

BRUCE A. BOLT² AND JAMES DORMAN

*Lamont Geological Observatory, Columbia University
Palisades, New York*

Abstract. Periods of spheroidal eigenvibrations, with order of spherical harmonic $n \geq 20$, have been computed for self-gravitating inhomogeneous spheres corresponding to a variety of earth models. The periods are used to deduce phase and group velocities for the fundamental and first higher modes of Rayleigh waves having periods less than 320 sec. The mathematical methods, program checks and estimations of numerical precision used in the work are presented in some detail. A comparison is made between phase and group velocities for different spherical models and with corresponding flat earth velocities calculated for the same physical parameters using the Thomson-Haskell matrix method for a nongravitating layered half-space. The comparison shows that the inclusion of gravity and sphericity increases the phase velocity by 0.25 km/sec (~ 5 per cent) near $T = 300$ sec and by 0.10 km/sec (~ 2.5 per cent) near $T = 150$ sec, where T is the period of the wave. In striking contrast, for $100 < T < 250$ sec, the group velocities for the spherical case remain within 1 per cent of the corresponding velocities for the horizontally layered case. A group velocity minimum occurs close to $T = 220$ sec. The variation with depth of the relative amplitudes of the radial and horizontal displacements and the perturbed gravitational potential is shown graphically for one earth model in the cases $n = 30, 80, 110$. On the basis of a comparison with these theoretical results, recent observations of Rayleigh wave phase and group velocities, as well as measured periods of some spheroidal oscillations, are consistent with an earth's mantle having a density distribution similar to that of model A constructed by Bullen, and having P and S velocity distributions similar to those calculated by Gutenberg. A group velocity dispersion curve which corresponds to the first higher mode (or first shear mode) of Rayleigh waves has also been computed for $20 < T < 350$ sec ($500 > n > 40$) for a continental model. Extrema of the curve which may be expected to produce relatively large amplitude arrivals occur at a velocity of 4.30 km/sec and period of 20 sec (a minimum) and at a velocity of 4.54 km/sec and period of 25 sec (a maximum). The period and velocity of the latter agree well with the period and velocity of the phase S_n described by Caloi and by Gutenberg and of the phase S_n of Press and Ewing.

Notation.

V Phase velocity
 U Group velocity
 T Period
 C Compressional-wave velocity
 S Shear-wave velocity
 ρ Density
 ω Angular frequency
 G Gravitational constant
 R Mean radius of the earth (6370 km)
 n Order of spherical harmonic
 m Spherical harmonic of order n
 g Acceleration of gravity

U Radial factor in the radial component of displacement
 V Radial factor in the horizontal component of displacement
 P Radial factor in the gravitational potential term
 ϵ Integration parameter
 λ Wavelength

Lamont Geological Observatory Contribution

on leave, Department of Applied Mathematics,
The University of Sydney, Australia.

Introduction. In a recent paper (hereafter called paper I), Dorman, Ewing, and Oliver [1960] presented computations of Rayleigh wave dispersion for eleven models of continental and oceanic crust-mantle systems having horizontal layers. These computations were used to explain observed mantle Rayleigh wave dispersion for periods of less than 250 sec. These au-

thors noted that, beyond periods of 200 sec, systematic differences between the observed and computed group velocities probably arose because of the inadequacy of the horizontal layer representation for dealing with long waves on the spherical earth. Also, the effect of the gravitational field was not allowed for in the computations used in paper I. In the present paper, which gives tables of phase and group velocities of Rayleigh waves in spherical, self-gravitating earth models, these questions are resolved. The full treatment of sphericity and gravity makes it possible to consider much more precisely the geophysical implications of the very long period surface wave observations which are now available.

The method used to obtain the new Rayleigh wave velocities depends on the calculation of the periods of the free spheroidal modes of vibration of a gravitating, inhomogeneous, elastic sphere. A solution of this problem was recently outlined by *Alterman, Jarosch, and Pekeris* [1959] and is based on earlier analyses of the sphere problem by *Rayleigh* [1906], *Love* [1911], *Hoskins* [1920], and *Jeans* [1923], among others.

Since the isolation of the relevant eigen-vibrations requires the integration of a sixth-order differential equation, the demands on computing machine time are relatively heavy, so the computed free periods of vibration as well as the corresponding velocities of the traveling Rayleigh waves are tabulated here in detail. The computations are of interest because they allow inferences on the structure of the earth's mantle to be drawn from recent good-quality observations of Rayleigh wave phase velocities [*Brune, Nafe, and Alsop*, in press; *Brune*, 1961], and group velocities [*Ewing and Press*, 1954a, b], as well as from measurements of the periods of the free spheroidal vibrations themselves [*Alsop, Sutton, and Ewing*, 1961; *Benioff, Press, and Smith*, 1961; *Bogert*, 1961; *Ness, Harrison, and Slichter*, 1961].

An outline of the mathematical procedures as well as numerical tests used to check the computer program and to estimate the precision of the results is included to allow comparison with the results of other workers [cf. *Alterman, Jarosch, and Pekeris*, 1959; *Pekeris, Alterman, and Jarosch*, 1961]. The tables included here for Rayleigh wave velocities in an earth model

having flat layers were computed by a program based on the Thomson-Haskell [*Haskell*, 1930] matrix iteration theory which was used in paper I. A comparison between the two calculations shows that, for $150 < T < 300$ sec, the corresponding group velocities differ by less than 1 per cent although the phase velocities for a sphere are from 2.5 to 5 per cent higher than the corresponding velocities for a half-space. It follows that the main inferences on the velocity structure of the upper mantle made in paper I remain firm, since they were based on observed group velocities of mantle Rayleigh waves.

Computations have been carried through for spherical earth models with two very different density distributions and three different velocity distributions in the mantle. Thus, in contrast to paper I, the importance to mantle wave propagation of major changes in the density distribution, a question previously raised by several workers, is brought out explicitly. Apparently, sensitivity to changes in the various physical parameters is such that empirical phase and group velocities measured with an accuracy of 0.01 km/sec for $50 < T < 300$ sec provide important information on the structure of the upper mantle. In particular, the combination of the Gutenberg velocity distribution with the Bullen A density distribution is closely consistent with free oscillation and phase and group velocity data in the above period range.

The calculation of Rayleigh wave dispersion in a sphere has also been carried out for the first higher mode (Sezawa's M_2 mode or the shear mode) for $20 < T < 150$ sec. Evidently, for the first time, this work has shown for a particular earth model the existence of turning points in the group velocity curve, a minimum near $C = 4.30$ km/sec, $T = 60$ sec, and a maximum near $C = 4.55$ km/sec, $T = 100$ sec. The result partly confirms an earlier prediction of *Oliver and Ewing* [1958] that there would be 'a flattening of the group velocity curve and perhaps another minimum at a velocity of about 4.4–4.7 km/sec at roughly a period of 20 sec.'

Mathematical formulation. Consider a spherical earth in which P and S seismic velocities, density, and gravity are all functions of radius r only. We investigate only the classical axisymmetric eigenvibrations which have spherical coordinates, (r, θ, φ) displacement

v, w) and cubical dilatation defined by *Ervin* [1920]:

$$2 \operatorname{curl} \mathbf{1} = 0$$

$$2 \operatorname{curl}_{\theta} \mathbf{1} = -(z \partial S / \partial \varphi) / r \sin \theta$$

$$2 \operatorname{curl}_{\varphi} \mathbf{1} = (z \partial S / \partial \theta) / r$$

$$\Delta = \left(\frac{1}{r^2} \frac{d(r^2 u)}{dr} - n(n+1) \frac{v}{r} \right) S$$

$z = d(rv)/dr - u$, and $S = S_n(\theta, \varphi)$ is a spherical surface harmonic.

The vibrations disturb the gravitational field by local changes in Δ and by the wrinkling of boundaries and interfaces; thus we express the gravitational potential Φ as $\Phi = P(r, \varphi)$. For a solid the equations of motion are (Ervin, 1920, p. 12) two second-order differential equations in the independent variable r , and changes in Φ yield another second-order equation with P as dependent variable. Suppression of the detail which is available in *Alterman, Pekeris, and Pekeris* [1959] we reduce these equations to the following set of six first-order

$$\dot{y}_5 = 4\pi G \rho y_1 + y_6 \quad (5)$$

$$\begin{aligned} \dot{y}_6 = & -4\pi G \rho n(n+1) \frac{y_3}{r} \\ & + n(n+1) \frac{y_5}{r^2} - \frac{2y_6}{r} \end{aligned} \quad (6)$$

where $\gamma^2 = \beta^2/\alpha^2 < 1$ and ρ and g are the unperturbed values for density and gravity at radius r . The coefficients of the equations are functions of α, β, ρ, g , which are the empirical parameters most directly available. The new variables y_i are defined to be

$$\begin{aligned} y_1 &= U & y_2 &= \widehat{rr} & y_3 &= V \\ y_4 &= \widehat{r\theta} & y_5 &= P & y_6 &= \dot{P} - 4\pi G \rho U \end{aligned}$$

where $\widehat{rr}, \widehat{r\theta}$ are stress components and U, V , and P are related to displacement as in (9).

In the remainder of this formulation displacements are supposed restricted to the solid mantle ($r > 3470$ km) of the earth models. (To the order of accuracy of the present numerical work this is known to be true at least for $n \geq 18$.) Equations 1 to 6 are then integrated by the Runge-Kutta method through homogeneous shells of thickness equal to the integration step. The shells are chosen sufficiently thin to obtain the required accuracy of integration and to adequately represent the variation of α, β, ρ , and g with r . Across each shell boundary all the y_i are taken continuous. At the free surface, $y_2 = y_4 = 0$; and the equality of the internal and external gravitational potentials yields $y_6 = -R y_5 / (n+1)$.

The crux of the method used to isolate the eigenvibrations is the property of linearity of equations 1 to 6. For a fixed n , three integrations to the surface were performed commencing from a prespecified depth $d = \epsilon \lambda$ (λ is wavelength, $\epsilon > 1$), with initial values $y_1 = y_3 = y_6 = 0$ and three linearly independent sets of values of y_2, y_4, y_5 . (Considerable freedom was available for the three latter values, since floating point arithmetic allowing magnitudes between approximately 10^{-38} and 10^{38} was used). In general these integrations yielded three independent sets of values of the dependent variables at the free surface, namely, $y_i^{(j)}, j = 1, 2, 3$. Since the actual solution of (1) to (6) which satisfies the surface boundary conditions is a linear

$$\begin{aligned} -2(1 - 2\gamma^2) \frac{y_1}{r} + \frac{y_2}{\rho \alpha^2} \\ + n(n+1)(1 - 2\gamma^2) \frac{y_3}{r} \end{aligned} \quad (1)$$

$$-\rho\{\omega^2 r^2 + 4gr$$

$$4(3 - 4\gamma^2)\beta^2\} \frac{y_1}{r^2} - 4\gamma^2 \frac{y_2}{r}$$

$$n(n+1)\rho\{gr - 2(3 - 4\gamma^2)\beta^2\} \frac{y_3}{r^2}$$

$$n(n+1) \frac{y_4}{r} - \rho y_5 \quad (2)$$

$$\dot{y}_3 = -\frac{y_1}{r} + \frac{y_3}{r} + \frac{y_4}{\rho \beta^2} \quad (3)$$

$$\rho\{gr - 2(3 - 4\gamma^2)\beta^2\} \frac{y_1}{r^2} - (1 - 2\gamma^2) \frac{y_2}{r}$$

$$\rho\{\omega^2 r^2 - 2\beta^2[2n(n+1)(1 - \gamma^2) - 1]\} \frac{y_3}{r^2}$$

$$\frac{y_4}{r} - \rho \frac{y_5}{r} \quad (4)$$

combination of the solutions now available, we have at the free surface

$$\begin{bmatrix} y_1^{(1)} & y_1^{(2)} & y_1^{(3)} \\ y_2^{(1)} & - & - & - \\ y_3^{(1)} & - & - & - \\ y_4^{(1)} & - & - & - \\ y_5^{(1)} & - & - & - \\ y_6^{(1)} & - & - & - \end{bmatrix} \cdot \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} U \\ 0 \\ V \\ 0 \\ P \\ -(n+1)P/R \end{bmatrix} \quad (7)$$

where a , b , c are undetermined multipliers, provided that ω is the eigenfrequency $\bar{\omega}$. In practice we calculate

$$\det(\omega) = \begin{vmatrix} y_2^{(1)} & y_2^{(2)} & y_2^{(3)} \\ y_4^{(1)} & y_4^{(2)} & y_4^{(3)} \\ y_6^{(1)} + \frac{n+1}{R} y_6^{(1)} & y_6^{(2)} + \frac{n+1}{R} y_6^{(2)} & y_6^{(3)} + \frac{n+1}{R} y_6^{(3)} \end{vmatrix}$$

for a series of frequencies ω , $\omega + d\omega \dots$ starting below the frequency of the expected eigenfrequency. The search is stopped when $\det(\omega)$ changes sign, and the eigenfrequency $\bar{\omega}$ corresponding to $\det(\omega) = 0$ is then estimated by inverse interpolation (using Lagrange's four-point formula) from the computed values of $\det(\omega)$. The smallest root of $\det(\omega) = 0$ corresponds to the fundamental mode of vibration while larger roots correspond to the higher modes (overtones) for a particular n . This method of searching for the correct eigenfrequency is based on the assumption that for the particular n there exists a solution of (1) to (6) in which the magnitudes of U , V , and P converge to zero with depth.

The variation of U , V , and P with depth, which is also of interest, may be determined by integration of (1) to (6) downward after the eigenfrequency has been found. For this computation, upward integration of (1) to (6) using the known eigenfrequency provides a set of surface values of U , V , and P from which a , b , and c are determined by (7) under the normalizing condition that U is unity at the surface.

For example, the surface value of V is $ay_3^{(1)} + by_3^{(2)} + cy_3^{(3)}$ where

$$a = \begin{vmatrix} y_2^{(2)} & y_2^{(3)} \\ y_4^{(2)} & y_4^{(3)} \end{vmatrix} \bigg/ \begin{vmatrix} y_1^{(1)} & y_1^{(2)} & y_1^{(3)} \\ y_2^{(1)} & y_2^{(2)} & y_2^{(3)} \\ y_4^{(1)} & y_4^{(2)} & y_4^{(3)} \end{vmatrix}$$

and similarly for b and c .

Downward integration of (1) to (6) is performed, and the desired amplitude distribution of U , V , and P is obtained. Results of procedure are shown in Figure 1 for $n = 80$, and 110, according to the parameter case 3. The characteristic amplitude distribution of these solutions throughout the mantle identify them as waves of the Rayleigh type [cf. Dorman and Prentiss, 1960, Figs. 4-7].

Since U , V , and P in the above discussion

radial factors, the vertical and horizontal displacements, u and v , and the gravitational disturbance, p , can be given as

$$\left. \begin{aligned} u &= U(r) S_n(\theta, \varphi) e^{i\bar{\omega}t} \\ v &= V(r) (\partial S_n(\theta, \varphi) / \partial \theta) e^{i\bar{\omega}t} \\ p &= P(r) S_n(\theta, \varphi) e^{i\bar{\omega}t} \end{aligned} \right\}$$

corresponding to the displacements in order Rayleigh waves.

The ultimate justification of a particular choice of ϵ to define the starting value $\epsilon\lambda$ for a vibration with wavelength $\lambda = 2\pi / (n + 1/2)$ is that, after following the above procedure, the computed variations with depth of U , V , and P , all show close convergence to zero for a depth less than $\epsilon\lambda$. This is achieved in the present work as exemplified in Figure 1.

It is of interest to compare the above mathematical procedure with that followed by [1959], who has integrated a fourth-order differential equation to obtain Rayleigh wave dispersion curves for a heterogeneous half-space. Satô denoted the ratio of the amplitude of horizontal and vertical displacements at the free surface by a . The method involved a set

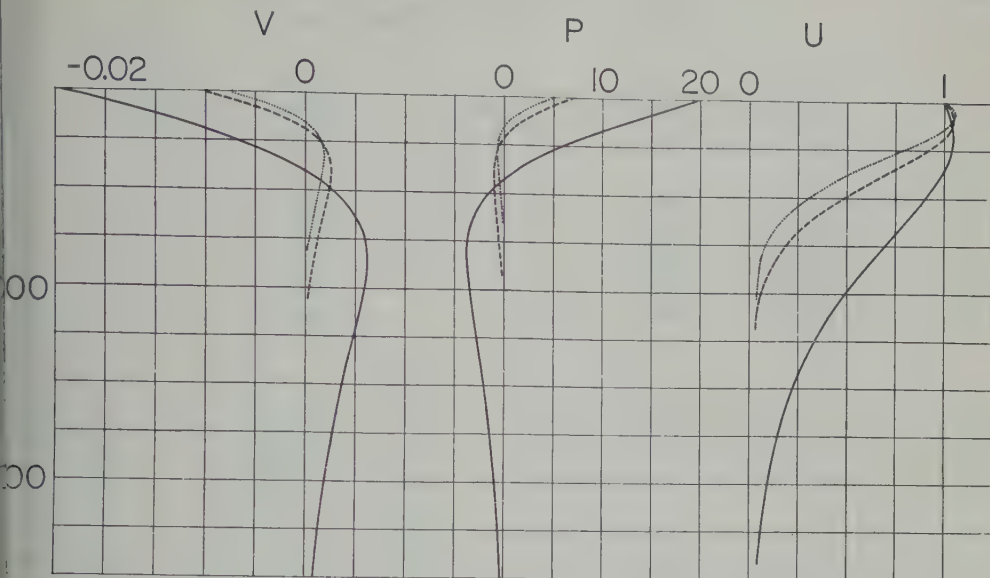


Fig. 1. Relative amplitudes of U , V , and P as functions of depth for: $n = 30$, solid curves; $n = 80$, dashed curves; $n = 110$, dotted curves. See equations 9.

a - c plane for that point which yields a tangent solution for the displacements with r . Thus in contrast with the method used by Satô's method involved integrations downwards as a means of locating the desired solution. The determination of the periods of free spheroidal waves now allows the calculation of phase velocities of Rayleigh-type waves. The appropriate formula is

$$c = \bar{\omega}R/(n + 1/2) \quad (10)$$

This formula has lately been studied in some detail by Matumoto and Satô [1954] and by Nafe and Alsop [in press]. From the theoretical and observational analyses of these waves, (10) is valid for $n \geq 20$ with an accuracy well within the numerical precision used in the present work. Since $C = c - \lambda dc/d\lambda$, the group velocity can be expressed conveniently using (10), as

$$C = c + (n + 1/2) dc/dn \quad (11)$$

For purposes of this paper the earth's mantle. For purposes of this paper the earth is considered to consist of concentrically arranged spherical shells. To simplify the integration problem, all shells are assumed to have equal thickness in a given computation. Each shell is a homogeneous material hav-

ing compressional velocity, shear velocity, and density which may be chosen arbitrarily to represent a stepwise approximation of any desired laws of variation as a function of r . In each case the parameters assigned to a shell are those given by the particular laws involved at the depth of the middle of the shell. As shown below in the section on numerical procedures, the method of computation produces a result which, strictly speaking, applies to the earth model made up of homogeneous shells. However, it will also be shown that the present results are indistinguishable in the given precision from results that would be obtained for earth models with continuously varying parameters.

Seven models of the earth's mantle, designated as cases 1 to 7, have been constructed for the research reported in this paper. The sources of the density and velocity distributions as well as certain parameters of each case are given in Table 1. In Tables 2 to 8 detailed data for each case are given. At the top of each table the parameters of the crustal shells and the top shell of the mantle are given. The interested reader may obtain the parameters of the remaining shells by interpolating the original data from the sources indicated in Table 1. Case 3 is virtually the same physical model as the 'Gutenberg' case of Pekeris, Alterman, and

TABLE 1. Definitions of the Cases Treated

Case	Density Distribution	Velocity Distribution	Presence of Gravity	Spherical Program	Shell Thickness, km	
1	Bullen A	Jeffreys	Yes	Yes	20	1
2	Bullen B	Gutenberg	Yes	Yes	20	1
3	Bullen A	Gutenberg	Yes	Yes	20	1
4	Bullen A	Gutenberg	No	No	20	
5	Bullen B	Gutenberg	No	No	20	
6	Bullen A	8099	Yes	Yes	10	2
7	Bullen A	Gutenberg	Yes	Yes	10	2

Jarosch [1961]. The calculated periods and phase velocities are also tabulated according to the order number of the oscillation in Tables 2 to 8.

Since the values of the physical parameters for each shell are constant throughout the shell and are obtained by interpolating from graphs, a different shell thickness implies a slightly different model. For this reason comparisons between cases in this paper are confined to those with the same shell thickness. Cases 4 and 5 have parameters identical to those of cases 3 and 2, respectively, but the computations are made with the program for the flat, nongravitating earth. Minor changes in the physical parameters used in the assumed 40-km con-

tinental crust, together with a change in thickness, distinguish case 3 from case 7. Case 7, with the smaller shell thickness, is used for short-period computations.

Either the *P* or *S* velocities derived by Jeffreys [1939] or those derived by Gutenberg [quoted by Bullard, 1957] are used in all cases except case 6, which is constructed to represent at least for *T* > 100 sec, the probable effect of the properties of the upper mantle beneath the oceans. Case 6 is identical below 11 km to the structure called case 8099 and defined in detail in paper I. The shear velocity in case 8099 is a low-velocity channel between depths of 0 and 220 km and joins smoothly onto Jeffreys' curve for *S* velocities near 350-km depth. The *P* velocity is based on the Jeffreys compressional wave velocity throughout. Since the boundary conditions used in the spherical earth program do not properly allow the inclusion of a liquid shell to represent the ocean water, a physically implausible crustal structure consisting of a single 10 km solid layer has been used purely as a mathematical device to simulate the oceanic crust. The properties of this layer were obtained by averaging the properties of the water with the solid crust above 10-km depth. In this case we may expect that the periods calculated at long wavelengths will be valid results for case 8099 with spherical curvature and gravity, because of the averaging effect of long waves.

Other important *P* and *S* velocity distributions derived from careful studies of seismic travel times are available [e.g. Lehmann, 1936] but the two cases discussed above are sufficient for a fruitful comparison between velocities of Rayleigh waves in earth models with and without a low-velocity channel. As is shown graphically in Figure 2, Jeffreys' solution for

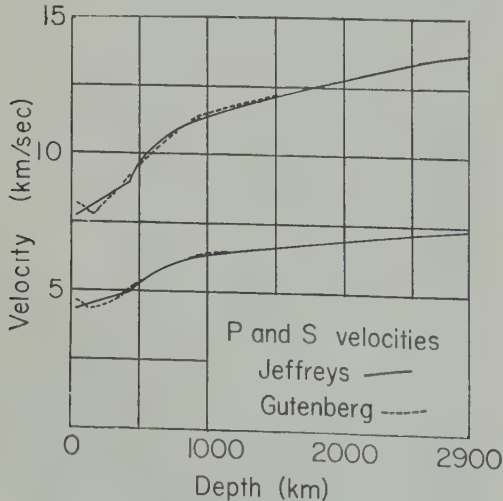


Fig. 2. *P* and *S* wave velocity distributions in the mantle used for computations of free periods. Solid curves are velocities of Jeffreys; dashed curves are velocities of Gutenberg. Crustal velocities used in each case are specified in the relevant table.

TABLE 2. Case 1

Parameters in the Crust			
α , km/sec	β , km/sec	ρ , g/cm ³	g , cm/sec ²
6.14	3.55	2.75	983.0
6.58	3.80	2.75	984.2
7.76	4.37	3.34	986.0

Calculated Periods and Velocities
Fundamentals

T , sec	c , km/sec	C , km/sec
322.6	5.51	
303.4	5.39	
286.4	5.27	3.85
271.6	5.17	3.78
258.4	5.08	3.76
246.5	5.00	3.73
235.7	4.92	3.71
225.8	4.86	3.72
216.7	4.80	3.71
208.4	4.74	3.71
200.6	4.69	3.73
193.4	4.65	3.73
186.7	4.61	3.74
180.3	4.58	3.75
174.4	4.54	3.76
168.9	4.51	3.77
163.7	4.49	3.78
158.8	4.46	3.79
154.1	4.44	3.80
149.8	4.42	3.80
145.6	4.40	3.81
141.7	4.38	3.82
138.0	4.36	3.82
134.4	4.35	3.83
131.0	4.33	3.83
127.8	4.32	3.83
124.8	4.31	3.84
121.9	4.29	3.84
119.1	4.28	3.85
116.4	4.27	3.85
113.9	4.26	3.85
111.4	4.25	3.85
109.1	4.24	3.85
106.8	4.23	3.85
104.7	4.22	3.85
102.6	4.22	3.85
100.6	4.21	3.85
98.7	4.20	3.85
96.9	4.19	3.86
95.1	4.19	3.85
93.4	4.18	3.85
91.7	4.18	3.85
90.2	4.17	3.85
88.6	4.16	3.85
87.1	4.16	3.85
85.7	4.15	3.84

TABLE 2. Continued

Calculated Periods and Velocities Fundamentals			
Order, n	T , sec	c , km/sec	C , km/sec
114	84.3	4.15	3.84
116	83.0	4.14	3.84
118	81.7	4.14	3.84
120	80.4	4.13	3.84
124	78.0	4.12	3.84
128	75.7	4.11	3.82
132	73.6	4.10	3.83
136	71.6	4.10	3.83
140	69.7	4.09	

TABLE 3. Case 2

Parameters in the Crust				
Thickness, km	α , km/sec	β , km/sec	ρ , g/cm ³	g , cm/sec ²
20	6.14	3.55	2.75	982.5
20	6.58	3.80	2.90	984.0
20	8.20	4.65	3.33	985.0

Calculated Periods and Velocities
Fundamentals

Order, n	T , sec	c , km/sec	C , km/sec
24	311.1	5.25	
26	293.2	5.15	
28	277.6	5.06	3.82
30	263.7	4.98	3.80
32	251.1	4.90	3.76
34	239.9	4.84	3.74
36	229.6	4.78	3.73
38	220.2	4.72	3.72
40	211.5	4.67	3.71
42	203.5	4.63	3.72
44	196.1	4.59	3.71
46	189.2	4.55	3.72
48	182.8	4.51	3.73
50	176.8	4.48	3.73
52	171.1	4.45	3.73
54	165.8	4.43	3.74
56	160.8	4.40	3.75
58	156.1	4.38	3.76
60	151.7	4.36	3.77
62	147.5	4.34	3.77
64	143.5	4.33	3.78
66	139.7	4.31	3.79
68	136.1	4.29	3.80
70	132.6	4.28	3.81

TABLE 3. Continued

Calculated Periods and Velocities Fundamentals			
Order, <i>n</i>	<i>T</i> , sec	<i>c</i> , km/sec	<i>C</i> , km/sec
72	129.4	4.27	3.82
74	126.3	4.26	3.82
76	123.3	4.24	3.83
78	120.4	4.23	3.84
80	117.7	4.22	3.85
82	115.1	4.21	3.85
84	112.6	4.21	3.86
86	110.2	4.20	3.87
88	107.9	4.19	3.87
90	105.7	4.18	3.88
92	103.6	4.18	3.89
94	101.5	4.17	3.89
96	99.6	4.17	3.89
98	97.7	4.16	3.90
100	95.9	4.16	3.90
102	94.1	4.15	3.91
104	92.4	4.15	3.91
106	90.8	4.14	3.92
108	89.2	4.14	3.92
110	87.6	4.13	3.92
112	86.2	4.13	3.93
114	84.7	4.13	3.93
116	83.3	4.12	3.93
118	82.0	4.12	3.94
120	80.7	4.12	3.94
124	78.2	4.11	3.95
128	75.9	4.11	3.93
132	73.7	4.10	3.95
136	71.6	4.10	3.96
140	69.6	4.09	

TABLE 4. Case 3

Parameters in the Crust				
Thickness, km	α , km/sec	β , km/sec	ρ , g/cm ³	g , cm/sec ²
20	6.14	3.55	2.75	983.0
20	6.58	3.80	2.90	984.2
20	8.20	4.65	3.34	986.0
Calculated Periods and Velocities Fundamentals				
Order, <i>n</i>	<i>T</i> , sec	<i>c</i> , km/sec	<i>C</i> , km/sec	
20	347.1	5.62		
22	324.4	5.48		
24	305.2	5.35		
26	288.5	5.24	3.76	

TABLE 4. Continued

Calculated Periods and Velocities Fundamentals			
Order, <i>n</i>	<i>T</i> , sec	<i>c</i> , km/sec	<i>C</i> , km/sec
28	273.7	5.13	3.69
30	260.7	5.03	3.63
32	248.9	4.95	3.61
34	238.2	4.87	3.60
36	228.5	4.80	3.58
38	219.5	4.74	3.57
40	211.2	4.68	3.58
42	203.6	4.63	3.58
44	196.4	4.58	3.58
46	189.7	4.54	3.59
48	183.5	4.50	3.60
50	177.6	4.46	3.61
52	172.1	4.43	3.62
54	166.9	4.40	3.63
56	162.0	4.37	3.65
58	157.3	4.35	3.66
60	152.9	4.33	3.66
62	148.8	4.31	3.68
64	144.8	4.29	3.69
66	141.0	4.27	3.70
68	137.4	4.25	3.71
70	134.0	4.24	3.72
72	130.8	4.22	3.73
74	127.7	4.21	3.74
76	124.7	4.20	3.75
78	121.8	4.19	3.75
80	119.1	4.17	3.76
82	116.5	4.16	3.77
84	114.0	4.16	3.78
86	111.6	4.15	3.79
88	109.3	4.14	3.79
90	107.1	4.13	3.80
92	104.9	4.12	3.81
94	102.9	4.12	3.81
96	100.9	4.11	3.82
98	98.9	4.11	3.83
100	97.1	4.10	3.83
102	95.4	4.09	3.83
104	93.7	4.09	3.84
106	92.0	4.08	3.84
108	90.4	4.08	3.85
110	88.9	4.08	3.85
112	87.4	4.07	3.86
114	86.0	4.07	3.86
116	84.5	4.06	3.86
118	83.2	4.06	
120	81.9	4.06	
140	70.6	4.03	
145	68.3	4.03	
150	66.1	4.02	
155	64.0	4.02	
160	62.1	4.02	
165	60.3	4.01	
170	58.6	4.01	

TABLE 5. Case 4

Parameters in the Crust			
Thickness, km	α , km/sec	β , km/sec	ρ , g/cm ³
20	6.14	3.55	2.75
20	6.58	3.80	2.90
20	8.20	4.65	3.34
Calculated Periods and Velocities Fundamentals			
T , sec	c , km/sec	C , km/sec	
300	5.06	3.96	
280	4.96	3.81	
260	4.85	3.69	
240	4.73	3.61	
220	4.61	3.57	
200	4.49	3.56	
180	4.38	3.58	
160	4.27	3.62	
140	4.18	3.68	
120	4.11	3.74	
100	4.05	3.80	
80	4.00	3.86	

TABLE 6. Case 5

Parameters in the Crust			
Thickness, km	α , km/sec	β , km/sec	ρ , g/cm ³
20	6.14	3.55	2.75
20	6.58	3.80	2.90
20	8.20	4.65	3.33
Calculated Periods and Velocities Fundamentals			
T , sec	c , km/sec	C , km/sec	
300	4.95	3.98	
280	4.87	3.88	
260	4.78	3.80	
240	4.68	3.74	
220	4.58	3.70	
200	4.48	3.69	
180	4.39	3.70	
160	4.30	3.72	
140	4.22	3.77	
120	4.16	3.82	
100	4.10	3.87	
80	4.06	3.92	

TABLE 7. Case 6

Parameters in the Crust				
Thickness, km	α , km/sec	β , km/sec	ρ , g/cm ³	g , cm/sec ²
10	4.00	1.50	1.90	983.0
10	7.82	4.61	3.34	983.0
10	7.82	4.61	3.34	983.0
10	7.82	4.61	3.34	983.0
10	7.82	4.61	3.34	983.0
Calculated Periods and Velocities Fundamentals				
Order, n	T , sec	c , km/sec	C , km/sec	
62	148.1	4.33		*
66	140.5	4.28		
70	133.6	4.25		3.67
74	127.4	4.22		3.68
78	121.7	4.19		3.69
82	116.4	4.17		3.70
86	111.6	4.15		3.71
90	107.2	4.13		3.73
94	103.0	4.11		3.73
102	95.7	4.08		3.75
106	92.4	4.07		3.76
110	89.3	4.06		3.77
114	86.3	4.05		3.78
118	83.6	4.04		3.78
122	81.1	4.03		3.79
126	78.6	4.02		3.80
130	76.4	4.02		3.80
138	72.2	4.00		

TABLE 8. Case 7

Parameters in the Crust				
Thickness, km	α , km/sec	β , km/sec	ρ , g/cm ³	g , cm/sec ²
10	5.89	3.40	2.00	983.0
10	6.24	3.60	2.75	983.0
10	6.41	3.70	2.75	983.0
10	6.58	3.80	3.00	983.0
10	8.04	4.63	3.33	983.0
Calculated Periods and Velocities First higher mode				
Order, n	T , sec	c , km/sec	C , km/sec	
40	148.0	6.68		5.27
50	124.3	6.38		
75	91.2	5.81		
100	73.0	5.45		4.32
125	61.1	5.22		4.29

TABLE 8. Continued

Order, <i>n</i>	Calculated Periods and Velocities First higher mode		
	<i>T</i> , sec	<i>c</i> , km/sec	<i>C</i> , km/sec
150	52.5	5.07	4.31
175	46.0	4.96	4.34
200	40.8	4.89	4.36
225	36.8	4.83	4.40
250	33.4	4.79	4.45
275	30.5	4.76	4.52
300	28.1	4.75	4.54
325	26.0	4.73	4.54
350	24.2	4.72	
375	22.7	4.69	4.53
400	21.3	4.69	
425	20.1	4.68	
450	19.1	4.66	4.46
475	18.1	4.65	
500	17.2	4.65	

velocity gradient has a discontinuity at a depth of 413 km (the 20° discontinuity); in contrast, Gutenberg's solution has the velocities first decreasing below about 50 km with a rapid increase at about 200 km. In the latter solution there is then an almost constant gradient to about 900 km, where there is an abrupt reduction of the gradient not found in the Jeffreys solution. It should be noted that despite the marked differences in the gradients the actual velocities agree within 0.3 km/sec at practically all depths.

Two principal density distributions in the mantle below the crust have been used. These are model A [Bullen, 1940, 1942] and model B [Bullen, 1950], based on different assumptions to represent conditions in the earth's interior. They are shown graphically in Figure 3. The density values of model A were computed from the Adams-Williamson equation and Jeffreys' velocities in the ranges of depth 33 to 413 km (region B), 1000 to 2900 km (region D) and 2900 to 4980 km (region E). In model B, the Adams-Williamson equation was used only for depths from 1000 to 2700 km (region D') and for region E; the density gradient between depths of 100 and 1000 km is fairly close to that in region D', a result which removes the significance of the transition of region C (413 to 1000 km).

The gravity distributions in the mantle used with model A and model B density distributions

are those already computed by Bullen for these models. For quite extreme density distributions in the earth, gravity values are not likely to deviate from 1000 cm/sec² by more than 1 per cent anywhere in the mantle [Bolt, 1957].

Numerical procedures and verification of the results. The problem was coded in Fortran for an IBM 7090 electronic computer using the formulation previously discussed for computation in floating point arithmetic (about eight decimal places). In the final program, called Sphos 2, the fourth-order method of Runge and Kutta is used, with fixed step size to integrate equations 1 to 6. Since this method is self-starting it is easy to apply. However, its relative slowness for continuing the solution is a disadvantage, most of the running time being used by four separate references per integration step to a subroutine, named Daux, for computing the six derivatives by equations 1 to 6. Since faster methods are available for continuing the solution, a fast-running program should contain separate programs for starting and for continuing the solution. Thus at the expense of greater programming and code checking effort a considerable saving in machine time can be achieved in subsequent work.

It was mentioned above that the solution strictly apply to an earth model which consists of thin, concentric, homogeneous shells rather than one with a continuous variation in physical

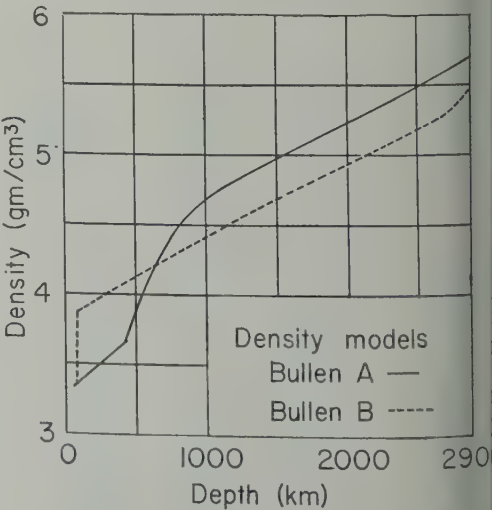


Fig. 3. Mantle density distributions of Bullen used for computations of free periods. Solid curve, model A; dashed curve, model B.

properties. We recall that the physical parameters assigned to each shell are the values at the depth of the center of the shell, estimated by careful graphical interpolation from continuous distributions such as those in Figures 2 and 3. The thickness of the shells is equal to the integration step used. Four references to the subroutines Daux occur at each integration step: one at the inner radius of the shell, two at the outer radius of the shell, and one at the outer radius of the shell. The physical properties used in each reference are the constant parameters for the corresponding shell. The right-hand sides of equations 1 to 6, which are evaluated by Daux, contain only the parameters themselves and not their derivatives with respect to r . Hence the integration effectively represents a solution of the differential equations through a stepwise variation of the physical properties. Similarly, the matrix iteration method for layers as used in paper I and in this paper provides a solution for a stepwise discontinuous distribution of properties. It was found in paper I that the layer thicknesses used there, usually 100 or 200 km in the deeper parts of the mantle and thinner above, were small enough to give results not significantly different from those which would be obtained from considering a continuous variation of parameters with depth. In the present computations, shell thickness is usually 20 km throughout the mantle, sufficiently thin so that results can be regarded as approximating to continuous distributions of parameters. For instance, in case 3, recomputation of eigenperiods using shell thickness 10 km rather than 20 km gave results differing by less than 0.1 sec from those in Table 4. The small differences occurring with the increase in step size are probably due to an increase in the truncation errors of the integration formula, but in any case they are within the precision required for the present study. The final program was tested in several ways. First, the period of oscillation for $n = 2$ for a homogeneous solid earth model (called model 1 by Alterman, Jarosch, and Pekeris [1959]) was determined, using 63 shells each 100 km thick, to be 44.37 minutes compared with 44.3 minutes given by Alterman, Jarosch, and Pekeris. Second, the greatest period of purely Rayleigh oscillations ($n = 0$) of a homogeneous earth model was computed by two inde-

pendent methods [Bolt, 1960]. Sphos 2 yielded a period of 8.73 minutes, when shells 100 km thick were used, in close agreement with a value of 8.75 minutes independently calculated by hand from the exact analytical solution. A further check became possible against values of periods and velocities corresponding to some of those for case 3 made available by Professor Pekeris. These values, now published [the 'Gutenberg' case, Pekeris, Alterman, and Jarosch, 1961], agree to three significant figures with the values in Table 4.

Tests of a range of values of ϵ , which defines the starting depth of the integration for each n , were made in order to select a value sufficiently large so that the free period is insensitive to moderate changes in ϵ but not so large as to make the integration needlessly long. By this criterion the value of ϵ selected for computation of the free periods in each case is sufficiently large so that the accuracy of the computed periods and velocities cannot be improved by increasing the magnitude of ϵ listed in Table 1. This conclusion is verified by the form of the relative amplitudes of the disturbance computed after the isolation of the eigenfrequencies. As in Figure 1, the amplitudes U , V , and P approach zero asymptotically in the vicinity of the depth $\epsilon\lambda$.

An interesting check on the Rayleigh wave velocities yielded by the present method is available at short periods by comparison with results for a flat-layered, nongravitating earth model calculated by the matrix method. Such a comparison, made graphically in Figure 5 for cases 3 and 4, shows that for $T = 25$ sec the two sets of values become almost identical. The Thomson-Haskell matrix formulation used to obtain the velocities in cases 4 and 5 (Tables 5 and 6) is quite unlike the approach using differential equations; the period equation is found from products of 4×4 matrices, one per layer, each of which is a function of c , T and the seismic parameter α , β , and ρ . Although the matrix formulation gives in general an efficient iterative method, the complete effect of sphericity and gravity on the period equation has not yet, to the authors' knowledge, been successfully treated in this way. As in Sphos 2, the program used for the layered half-space calculations requires that the physical parameters used in each layer be sampled from the

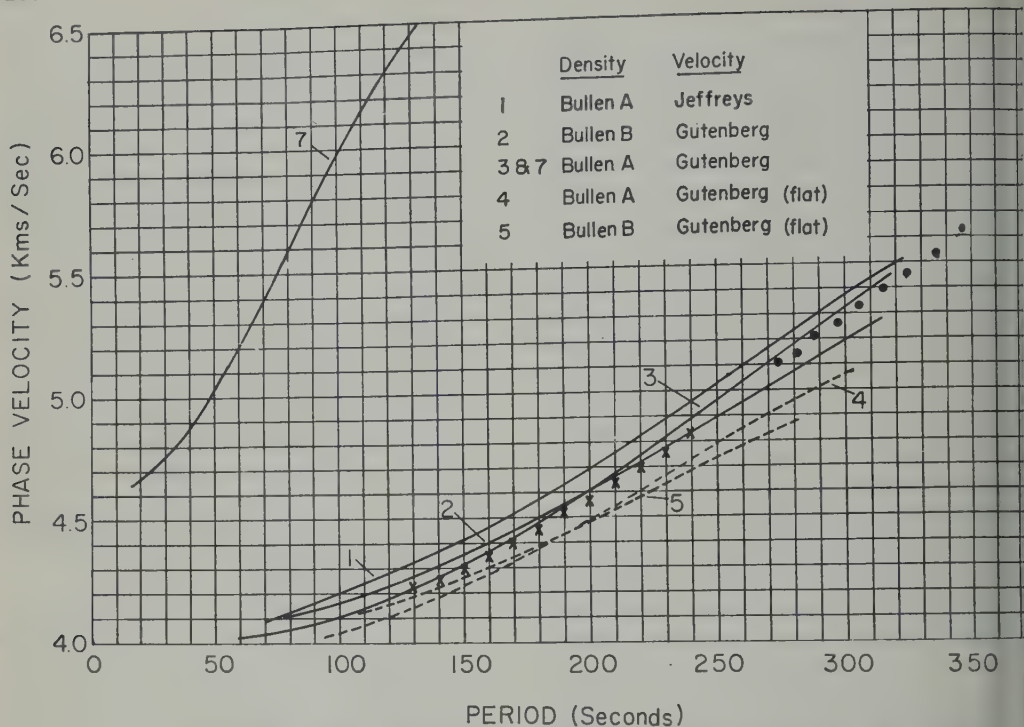


Fig. 4. Theoretically computed and experimentally observed phase velocity data for mantle Rayleigh waves. Solid curves are computed by the program for spherical, gravitating earth (Sphos 2); dashed curves by the program for the flat, layered earth. Case numbers are indicated. Cases 3 and 4 are based on the same model as cases 2 and 5. X's indicate mean empirical velocities given by *Brune* [1961]. Solid circles indicate mean empirical velocities given by *Brune, Nafe, and Alsop* [in press]. Case 3, from the Gutenberg velocity and Bullen A density models, is the most satisfactory mantle structure.

empirical curves at the depth of the midpoint of the layer. Another approach to the Rayleigh wave propagation in a sphere has been followed by *Yanovskaya* [1958], who, by using asymptotic expansions of the Hankel functions, obtained corrections $O(\lambda/2\pi R)$ to the terms of the determinantal period equation for a single solid layer over a half-space. The formulation is, however, rather more complicated than the one in terms of the differential equation used here.

Geophysical inferences. The periods and phase and group velocities tabulated in Tables 2 to 8 allow a number of important geophysical conclusions to be drawn. We see from Figure 4, where phase velocity curves are drawn for five cases, that the combined effect of sphericity and gravitation of the earth is to produce a marked increase of phase velocity with wavelength with respect to flat earth values. Cases 4 and 5, drawn in Figure 4 with broken lines,

were computed by the Thomson-Haskell matrix method for flat layers. They may be compared directly with cases 3 and 2 calculated using Sphos 2.

In both comparisons the total effect of spherical curvature and gravity is of the same order of magnitude, an increase in phase velocity of 0.1 km/sec (~ 5 per cent) at $T = 300$ sec, about 0.10 km/sec (~ 2.5 per cent) at $T = 250$ sec; even down to $T = 100$ sec the phase velocities in a spherical earth model are about 0.1 km/sec above those for a flat one. This is notable since some results on crust and upper mantle structure based on the correlation between observations of long-period waves and computed curves for a flat earth model have been published. To clear up this point we have therefore carried the calculations for a modified case 3 (case 7) down to periods near 25 sec ($n = 450$) using both the spherical and

programs. Case 7 differs from case 3 by having a decreased step thickness of 10 km to permit calculation at the shorter wavelengths involved and by having a 40-km continental crust as specified in the legend of Figure 5. The spherical earth case used in Figure 5 for comparison with case 7 has the same differences with respect to case 4.

The conclusion from the curves in Figure 5 is that at least for the fundamental mode, the effect of sphericity and gravity is negligible for the purpose mentioned only up to the period range from 30 to 40 sec. Obviously, the desired accuracy in an experimental study must be checked carefully before selecting a flat earth dispersion calculation for phase velocity.

To derive from the computed phase velocities the following approximate empirical relationship between the wavelength, the phase velocity in a gravitating sphere, and the corresponding velocity c_h for a layered half-space with no gravity.

$$c \approx c_h(1 + 0.00016T) \quad (12)$$

This equation allows the estimation of c within 1 per cent for $100 < T < 300$ sec from values computed for a flat earth model.

It is convenient here to compare in the same figure the group velocities for the cases discussed above. This is done graphically for cases 2 and 3 in Figure 6. Whereas the phase velocities are significantly different for $100 < T < 250$ sec, the group velocities in both cases 3 and 4 and cases 2 and 5 remain equal within 1 per cent. In view of (11) in its earlier form, this result agrees with that

$$(c - c_h)/\lambda \approx d(c - c_h)/d\lambda \quad (13)$$

It agrees with (12). It follows that the main conclusions reached in paper I from a comparison between observed group velocity data and theoretical group velocity dispersion computations for horizontal layers remain largely valid. Figure 6 also shows that cases 2 and 3 yield maximum group velocities of 3.71 km/sec at $T = 20$ sec and 3.57 km/sec at $T = 220$ sec, respectively. These periods are remarkably near the estimate of about 4 minutes made earlier by Ewing [1937], and the curve of case 3 agrees well with a more quantitative estimate made by Ewing and Press [1954a].

The spherical curvature also modifies the distribu-

tion of amplitudes of the disturbance as a function of depth. The three examples of amplitude distribution drawn in Figure 1 may be compared with particle amplitude profiles for Rayleigh waves in a heterogeneous, solid half-space [Dorman and Prentiss, 1960]. A comparison between displacements for the same phase velocity and similar physical parameters computed by the spherical and flat earth programs indicates that the effect of sphericity is to concentrate the motion nearer the surface.

The main differences between the theoretical Rayleigh wave phase velocity curves for the spherical earth models shown in Figure 4 are a consequence of the differences between the two density distributions and the two velocity distributions used. For periods $100 < T < 350$ sec the corresponding range of wavelengths is approximately $400 < \lambda < 2000$ km. The calculated relative displacements for orders $n = 30$ ($\lambda = 1312$ km), $n = 80$ ($\lambda = 497$ km), and $n = 110$ ($\lambda = 362$ km) displayed in Figure 1 demonstrate that waves of this length respond mostly to the physical properties of the upper mantle. As a consequence, the phase velocities of case 1 based on Jeffreys' seismic velocities are greater in the period range 100 to 350 sec than those of case 3, which is based on Gutenberg's seismic velocities. This is because case

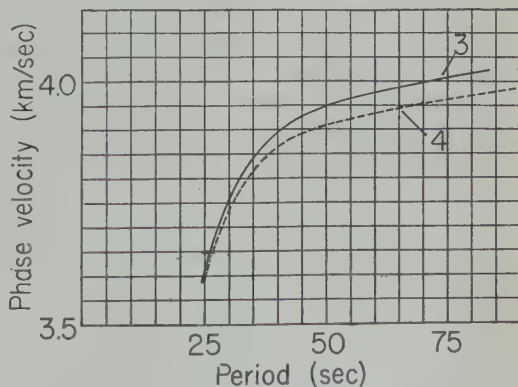


Fig. 5. Phase velocity dispersion curves from modifications of cases 3 and 4 for short-period computations. The shell thickness, 20 km in cases 3 and 4, has been changed to 10 km, and the crustal layers have been changed slightly for these computations. These curves show the difference between computed phase velocities for a spherical gravitating earth (case 3), and for a flat, non-gravitating earth (case 4) for comparable physical models.

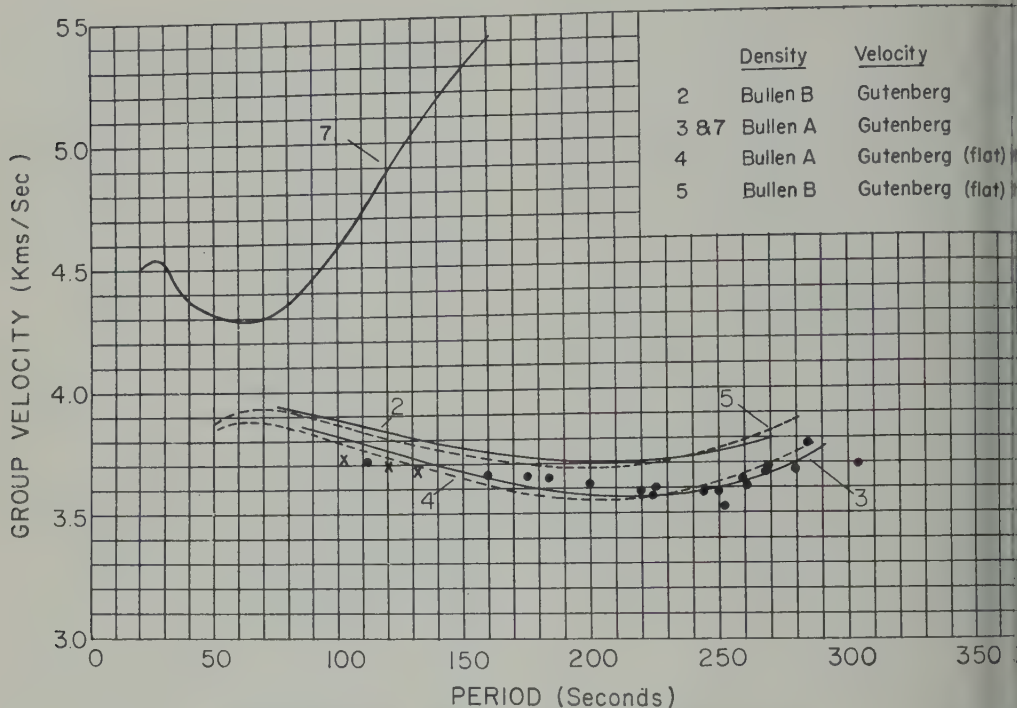


Fig. 6. Theoretically computed and experimentally observed group velocity data for mantle Rayleigh waves. Solid curves are computed for a spherical, gravitating earth; dashed curves for a flat, nongravitating earth. In contrast to phase velocity results, flat and round earth group velocity curves for similar physical models differ by only minor amounts. X's indicate group velocity observations from Assam earthquake of August 15, 1950 [Ewing and Press, 1954a]. Solid circles indicate group velocity observations from Kamchatka earthquake of November 4, 1952 [Ewing and Press, 1954b]. The data in this form also favor case 3.

3 has a low-velocity channel down to a depth of about 300 km (cf. Fig. 2) and in spite of the fact that case 1 has a deficiency of velocity immediately below the Mohorovicic discontinuity. (It is known from paper I that very low velocities at very shallow depths in the mantle cause case 1 to have lower phase velocities than case 3 at periods less than 60 sec.)

Similarly, Figure 4 shows that the higher densities at depths less than 600 km in model B cause the phase velocities of case 2 to exceed those of case 3 for $T < 200$ sec ($\lambda < 900$ km); for $T > 200$ sec the phase velocities in case 2 become the smaller in response to the marked deficiency of density in model B with respect to model A at depths greater than 600 km. (This density difference remains near 0.3 g/cm^3 (~ 7 per cent) in the mantle below 1000 km (cf. Fig. 3).)

Also shown in Figure 4 are empirical phase

velocities [Brune, Nafe and Alsop, in press] from observations at Palisades of the Chile earthquake of May 22, 1960 (continental station of path approximately 60 per cent), mean phase velocities for typical continental great-circle paths approximately 40 per cent continental [Brune, 1961]. Brune estimates accuracy of the measured velocities to be within ± 0.5 per cent (personal communication). In general the observations closely follow the curve for case 3 in the range $120 < T < 330$ sec, falling less than 0.03 km/sec below it at 120 sec period. Therefore, an earth model having a mantle with a velocity distribution similar to Gutenberg's model (including a low velocity channel) and a density distribution similar to that of Bullen's model A is most consistent with the data. Taken together, the theoretical curves suggest that a modification of case 3 having a slightly wider low-velocity channel and slightly

densities below 600 km may give better agreement with the observations.

Group velocity observations from the Kamata shock of November 4, 1952, and the earthquake of August 15, 1950 [Ewing Press, 1954*a*, *b*], are plotted in Figure 6. The observations show some scatter, but in a way that is in good agreement with the phase velocity data clearly showing the mantle structure of case 3 over that of case 2. The group velocity curve for case 1 is also plotted in Figure 6, but examination of the theoretical data in Table 2 shows that it does not agree with the observations. A similar conclusion with regard to the same velocity and density distributions taken together was reached in paper I, in which the same group velocity dispersion data were used. Therefore, in the present case this result was expected since we have shown above that the effect of spherical curvature on group velocity dispersion is small in the period range under consideration.

Further evidence of the relative superiority of the mantle structure similar to that of case 3 is available from spectral analyses of Palisades programs (from both vertical and north-south component instruments) of the Chilean earthquake mentioned above [Alsop, Sutton, and Brune, 1961]. These analyses yielded estimated periods between 250 and 500 sec which were identified as corresponding to fundamental free spheroidal oscillations of the earth. A comparison between the theoretical periods of case 3 and harmonics of order 20 to 34 and these experimentally determined periods is made in Table 3. Standard errors for the various periods are given by Alsop, Sutton, and Ewing, but the agreement appears significant.

In a recent paper, Pekeris, Alterman, and Brune [1961] presented theoretical spheroidal free periods for $2 \leq n \leq 61$. They compare their theoretical free periods for two earth models with various reported observations and state that the same conclusion reached in this paper, namely, that their 'Gutenberg' case, which is identical with case 3, is the preferred

one. It should be pointed out that the Bullen A velocity distribution and the Gutenberg velocity distribution used in case 3 are not mutually consistent according to the method used by Bullen to derive the density distribution, since it involved a relationship to the seismic veloc-

TABLE 9. Comparison of Observed Free Periods in Seconds of Fundamental Spheroidal Oscillations with Theoretical Values for Case 3

Order, n	Theoretical	Observed	
		Z	N-S
20	347	347	
21	335	335	335
22	324	325	325
23	314	315	315
24	305	307	306
25	296	297	297
26	288	289	289
27	281	281	
28	274		
29	267	268	
30	261	262	261
31	255	255	
32	249	249	
33	243		
34	238	239	

ities of Jeffreys [Bullen, 1940]. Consideration of this relationship is beyond the scope of this paper, but, clearly, it provides an additional restriction which should be observed in a solution which is satisfactory from all points of view.

As mentioned above, case 6 (see Table 7) was modeled after case 8099 of paper I to represent the effect on long-period waves of the possible structure of the mantle beneath the oceans. The velocities listed in Table 7 may be compared with those for similar periods in case 8099 determined from the layered half-space calculation (paper I, p. 94). It is seen that for $T > 100$ sec the difference between the two sets of data is given approximately by (12), as expected. However, for $T < 100$ sec the difference becomes smaller with decreasing period and vanishes at about $T = 70$ sec, probably indicating difficulty with the top layer approximation of case 6. At $T > 100$ sec, at which the approximation appears reliable, this theoretical case gives phase velocities which exceed those of case 3, the most suitable continental case, by 0.03 km/sec or less. This agrees with the experimental data reported by Brune [1961], who finds that, for complete great-circle paths, phase velocities over more nearly oceanic paths are slightly higher than over more nearly continental paths. Case 6 agrees

quite well with the oceanic data of Brune in the region of overlap, $120 < T < 150$ sec. Since accurate spherical-earth theoretical dispersion data in the period range $T < 100$ sec are not presently available for oceanic structures, we cannot at the moment reconsider the solution for shallow suboceanic mantle structure given in paper I. However, present theoretical data for $T > 100$ sec is in agreement with the conclusions previously reached.

The previous discussion has dealt only with the fundamental Rayleigh mode. The computed Rayleigh wave velocities for the first higher mode in case 7 are also given in Figures 4 and 6. The group velocity dispersion curve is of particular interest. Its main features are a local minimum of velocity of 4.29 km/sec at $T = 60$ sec and a local maximum of velocity of 4.54 km/sec at about $T = 25$ sec. Since waves associated with these stationary points should possess relatively large amplitudes, these values, which are calculated for a continental crust, suggest that some of the waves called S_n by Caloi [1954] and Gutenberg [1955] and S_n (type I) by Press and Ewing [1955] may correspond to the local maximum of the higher-mode dispersion curve. According to Press and Ewing, S_n (type I) is 'characterized by large vertical motion consisting of one or two oscillations with periods of about 20-30 seconds. The average velocity of the phase is about 4.58 km/sec.' This description fits very well the waves which might be expected on the basis of the higher-mode curve computed here. Oliver and Ewing [1958] have suggested that some S_n or S_n waves might possibly be explained in terms of normal mode propagation in a continental crust. The theoretical curves available to these authors were based on a layered half-space calculation for a double-layered crust and a homogeneous mantle and extended only to periods of about 15 sec. The present extension of this theoretical curve to much longer periods, including the effects of curvature, gravity, and mantle heterogeneity, places the identification of the mechanism of these waves on a much firmer basis.

In a paper just published, Alterman, Jarosch, and Pekeris [1961] list theoretical phase and group velocities for three earth models, one of which is the 'Gutenberg' model cited above. The other two models have (1) densities of

Bullen's model B and the Jeffreys-Bullen velocities, and (2) densities of Bullen's model A and the Jeffreys-Bullen velocities, with some modifications introduced by Dorman, Ewing, and Oliver [1960]. For periods less than 150 sec the phase velocities are calculated using an approximation called the earth-flattening method, rather than the complete differential equation method used in the present paper. Alterman, Jarosch, and Pekeris find that for each model the effect of the core is negligible for $n > 1$, a result which is basic for this paper. They also conclude that, of the three earth models considered, the Gutenberg case, corresponding to case 3 of this paper, is the one which gives best agreement with the observed phase velocity data of Ewing and Press [1956] and Nafe and Brune [1960]. The calculations presented in Figure 5 agree well with their estimate of 150 sec as the longest period at which 1 per cent accuracy in the Rayleigh wave phase velocity can be obtained from a flat-earth calculation. The agreement between the numerical results of these two papers indicates that a high degree of confidence can be placed in the accuracy of the rather lengthy computations.

Acknowledgments. We are grateful for criticisms made by Maurice Ewing and Jack Oliver during the preparation of this paper. The computations were made possible by the provision of IBM 7090 computer time by the Theoretical Division of the National Aeronautics and Space Administration, and International Business Machines Corporation, Watson Scientific Computing Laboratory at Columbia University. Professor C. L. Pekeris of the Weizmann Institute, Israel, kindly made available some of his results on Rayleigh wave velocities prior to their publication.

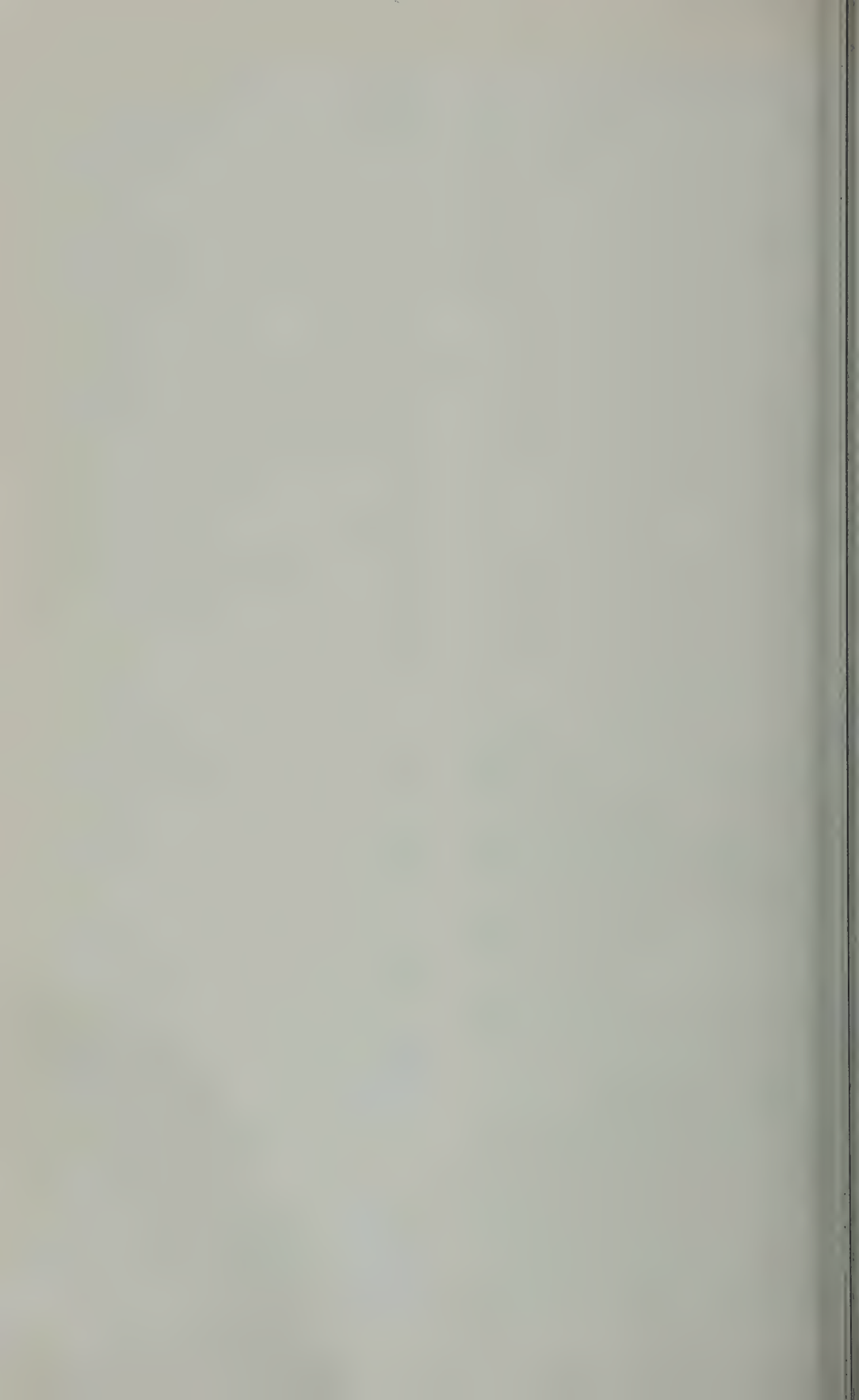
The research was supported in part by the National Science Foundation and the Air Force Cambridge Research Center. The first author was assisted by a Fulbright Research Scholarship.

REFERENCES

- Alsop, L. E., G. H. Sutton, and M. Ewing, Free oscillations of the earth observed on strain and pendulum seismographs, *J. Geophys. Research*, **66**, 631-641, 1961.
- Alterman, Z., H. Jarosch, and C. L. Pekeris, Oscillations of the earth, *Proc. Roy. Soc. London*, **252**, 80-95, 1959.
- Alterman, Z., H. Jarosch, and C. L. Pekeris, Propagation of Rayleigh waves in the earth, *Geophys. J.*, **4**, 219-241, 1961.
- Benioff, H., F. Press, and S. Smith, Excitation

- the free oscillations of the earth by earthquakes, *J. Geophys. Research*, **66**, 605-619, 1961.
- Hert, B., An observation of the free oscillations of the earth, *J. Geophys. Research*, **66**, 643-646, 1961.
- Hoskins, B. A., Earth models with chemically homogeneous cores, *Monthly Notices Roy. Astron. Soc., Geophys. Suppl.*, **7**, 372-378, 1957.
- Hoskins, B. A., Spheroidal oscillations of the moon, *Nature*, **188**, 1176-1177, 1960.
- Jeffreys, H., Radiation pattern of Rayleigh waves from the southeast Alaska earthquake of July 1958, Symposium on Earthquake Mechanism, *Publ. Dominion Observatory, Canada*, **10** (10) 1961.
- Jeffreys, H., J. N. N., J. E. Nafe, and L. E. Alsop, The phase shift of surface waves on a sphere, *Bull. Seism. Soc. Am.*, in press.
- Kanamori, H., The density within the earth, *Verhandl. Ned. Geol. Mijnbouw. Genoot. (Geol. Ser.)*, **18**, 23-41, 1957.
- Kanamori, H., The problem of the earth's density variation, *Bull. Seism. Soc. Am.*, **30**, 235-250, 1940.
- Kanamori, H., The density variation of the earth's central core, *Bull. Seism. Soc. Am.*, **32**, 19-29, 1942.
- Kanamori, H., An earth model based on a compressibility-pressure hypothesis, *Monthly Notices Roy. Astron. Soc., Geophys. Suppl.*, **6**, 1-59, 1950.
- Kanamori, H., P., L'Astenosfera come canale-guida dell'energia sismica, *Ann. geofis. Rome*, **7**, 491-501, 1954.
- Kanamori, H., J. M. Ewing, and J. Oliver, Study of surface-wave velocity distribution in the upper mantle from mantle Rayleigh waves, *Bull. Seism. Soc. Am.*, **50**, 87-115, 1960.
- Kanamori, H., and D. Prentiss, Particle amplitude profiles for Rayleigh waves on a heterogeneous earth, *J. Geophys. Research*, **65**, 3805-3816, 1960.
- Kanamori, H., W. S. Jardetzky, and F. Press, *Elastic Waves in Layered Media*, McGraw-Hill Book Co., New York, 1957.
- Kanamori, H., M., and F. Press, An investigation of mantle Rayleigh waves, *Bull. Seism. Soc. Am.*, **44**, 141-148, 1954a.
- Kanamori, H., M., and F. Press, Mantle Rayleigh waves from the Kamchatka earthquake of November 1952, *Bull. Seism. Soc. Am.*, **44**, 471-479, 1954b.
- Kanamori, H., M., and F. Press, Rayleigh wave dispersion in the period range 10 to 500 seconds, *Trans. Am. Geophys. Union*, **37**, 213-215, 1956.
- Kanamori, H., B., Channel waves in the earth's crust, *Geophysics*, **20**, 283-294, 1955.
- Kanamori, H., N. A., Dispersion of surface waves on multilayered media, *Bull. Seism. Soc. Am.*, **43**, 17-34, 1953.
- Hoskins, L. M., The strain of a gravitating sphere of variable density and elasticity, *Trans. Am. Math. Soc.*, **21**, 1-43, 1920.
- Jeans, J. H., The propagation of earthquake waves, *Proc. Roy. Soc. London A*, **102**, 554-574, 1923.
- Jeffreys, H., The times of *P*, *S* and *SKS* and the velocities of *P* and *S*, *Monthly Notices Roy. Astron. Soc., Geophys. Suppl.*, **4**, 498-533, 1939.
- Lehmann, I., Velocities of longitudinal waves in the upper part of the earth's mantle, *Ann. géophys.*, **15**, 93-118, 1959.
- Love, A. E. H., *Some Problems of Geodynamics*, Cambridge Univ. Press, Cambridge, 1926.
- Matumoto, T., and Y. Satô, On the vibration of an elastic globe with one layer. The vibration of the first class, *Bull. Earthquake Research Inst., Tokyo Univ.*, **32**, 247-258, 1954.
- Nafe, J. E., and J. N. Brune, Observations of phase velocity for Rayleigh waves in the period range 100 to 400 seconds, *Bull. Seism. Soc. Am.*, **50**, 427-439, 1960.
- Ness, N., J. Harrison, and L. Slichter, Observations of the free oscillations of the earth, *J. Geophys. Research*, **66**, 621-629, 1961.
- Oliver, J., and M. Ewing, Normal modes of continental surface waves, *Bull. Seism. Soc. Am.*, **48**, 33-49, 1958.
- Pekeris, C. L., Z. Alterman, and H. Jarosch, Comparison of theoretical with observed values of the periods of free oscillation of the Earth, *Proc. Natl. Acad. Sci. U. S.*, **47**, 91-98, 1961.
- Pekeris, C. L., and H. Jarosch, The free oscillations of the earth, in *Contributions in Geophysics*, vol. I, Pergamon Press, London, 1958.
- Press, F., and M. Ewing, Waves with *P_n* and *S_n* velocity at great distances, *Proc. Natl. Acad. Sci. U. S.*, **41**, 24-27, 1955.
- Rayleigh, Lord, On the dilatational stability of the earth, *Proc. Roy. Soc. London A*, **77**, 486-499, 1906.
- Satô, Y., Numerical integration of the equation of motion for surface waves in a medium with arbitrary variation of material constants, *Bull. Seism. Soc. Am.*, **49**, 57-77, 1959.
- Stoneley, R., Surface waves associated with the 20° discontinuity, *Monthly Notices Roy. Astron. Soc., Geophys. Suppl.*, **4**, 39-43, 1937.
- Yanovskaya, T. B., The dispersion of Rayleigh waves in a spherical layer, *Bull. Acad. Sciences USSR, Geophys. Ser.*, no. 7, 801-817, 1953.

(Manuscript received March 2, 1961; revised June 20, 1961.)



Steady-State Calculations Bearing on Geological Implications of a Phase-Transition Mohorovicic Discontinuity

G. W. WETHERILL¹

*Institute of Geophysics and Department of Geology
University of California, Los Angeles, California*

Abstract. The steady-state approximation is used to calculate the temperature as a function of depth within the earth's crust for different crustal models. It is assumed that the crust is bounded below by a phase transition, and the depth of this transition is calculated for several typical transitions. If experimental uncertainties are taken into consideration, there is no difficulty in fitting phase transitions which have been observed in the laboratory with depths of the continental Mohorovicic discontinuity, but it is difficult to obtain reasonable oceanic crustal thicknesses with the same phase transition. These calculations indicate that: (1) If the Mohorovicic discontinuity results from a phase transition similar to the albite + nepheline \rightarrow jadeite transition, the continental crust should have had an average elevation of over 4 km about 2000 million years ago. (2) An increase by a factor 1.25 in the continental heat flow can result in appreciable changes in elevation. The particular model used would have about a 2-km increase in elevation. (3) It is difficult to reconcile the absence of greater elevation and crustal thickening in the area of high heat flow in the southeastern Pacific Ocean with the phase-change hypothesis. (4) The mechanism of mountain building discussed by Kennedy and by MacDonald and Ness can produce appreciable crustal elevations if the time scale for the process is sufficiently long.

Introduction. It has been emphasized by King [1958], Kennedy [1959], and MacDonald and Ness [1960] that, if the Mohorovicic discontinuity (M) indeed represents a phase transition, this may have very important implications regarding such geological problems as the difference between continents and ocean basins, the permanence of continents, the elevation of plateaus, the formation of geosynclines, the origin of mountains. These implications stem from the fact that the pressure at which the phase change takes place will be temperature dependent and M will migrate as a result of a process which changes the temperature within the earth's crust. The resulting changes in crustal thickness will produce changes in elevation in accordance with the requirements of the equilibrium. Therefore the nature of M has widespread importance to many fields of geology, and, conversely, geological observation is expected to contribute information concerning the nature of M.

In general, the prediction of these effects is mathematically difficult. It is possible, however, to gain considerable insight into these problems with extremely elementary calculations by making use of the steady-state approximation. This insight permits these effects to be semiquantitatively considered without the necessity of carrying out formidable numerical calculations. In this paper some results of such calculations are presented. Because of uncertainties regarding the actual composition of the earth's crust, the properties of natural phase transitions, and the probability that the phase transitions take place over a range of pressures, these results are suggestive rather than definitive.

The steady-state or equilibrium approximation has also been made by MacDonald and Ness, but the results of the calculations in the present paper differ so markedly from theirs that these results are believed to be of interest. (A discussion of the origin of these differences is given in the Appendix.)

The relevant equation is

$$\frac{\partial T}{\partial t} = \frac{k}{\rho c} \nabla^2 T + \frac{\epsilon}{\rho c} + \mathbf{V} \cdot \nabla T \quad (1)$$

¹Part of this work was done while the author was on the staff of the Department of Terrestrial Magnetism, Carnegie Institution of Washington, while visiting at the California Institute of Technology.

where

- $T(x, y, z, t)$ = temperature.
 k = thermal conductivity.
 ρ = density.
 c = specific heat at constant volume.
 $\epsilon(x, y, z, t)$ = rate of heat production per unit volume = ρq where q is the heat production unit mass.
 $V(x, y, z, t)$ = velocity of the matter at any point.

The steady-state solution for constant ϵ is

$$T = T_a + \frac{Q_a}{k}(x - a) - \frac{\epsilon}{2k}(x - a)^2 \quad (2)$$

where x represents the depth in the crust, a is some depth (e.g. the surface) at which the temperature T_a and the heat flux Q_a are known.

Because this solution is one-dimensional, effects resulting from the curvature of the earth are neglected, and the assumption of a steady state requires that the parameters (ϵ , k , etc.) and the model employed have not changed appreciably in a time long compared with the characteristic thermal relaxation time ($d^2\rho c/k$) of a layer of depth d .

For the crust and upper mantle these times are short compared with the age of the earth and also compared with the time in which important changes occur due to decay of the long-lived natural radioactive nuclides.

Temperatures in a continental crust bounded below by typical phase transitions. The Clapeyron curves for the phase transitions will be represented by the equation $T = bp - c$, where T and p are the temperature and pressures and b and c are constants [Thompson, 1955]. In accordance with the assumptions of MacDonald and Ness, the following values of b and c are adopted:

Albite Decomposition	b , °C cm ² /kilobar	c , °C
Albite + nepheline → jadeite and albite → jadeite + quartz	0.0520	50 to 250
Sapphirine + enstatite → pyrope	0.105	500 to 1000

In view of the uncertainties in the experimental data, different assumptions regarding the phase transitions could be made, but the above values of b and c are used in order to make the results comparable with those of MacDonald and Ness.

In Figure 1 the Clapeyron curves defining the phase boundaries are plotted, together with the temperature distribution curves for single- and double-layer models. These models are shown in Figure 2. The values chosen for the heat production and conductivity are in the range given by Birch [1942, 1955]. The point of intersection of a Clapeyron curve and a temperature curve gives the temperature and pressure within the earth at which M would occur if it were a result of a phase transition characterized by the Clapeyron curve.

In the case of the pyrope transition, with $c = 500^\circ\text{C}$, it may be seen that for two of the single-layer models and both the two-layer models the phase transition takes place at depths similar to those of M . These models have in common the property that, if the phase transition is taken as marking the lower boundary of the crust, more than half of the radioactive activity will be in the crust. With $\epsilon/2k = 0.075^\circ\text{C}/\text{km}^2$ the crust contains so much radioactive activity that the crustal heat production equals the surface heat flow before the depth of the phase transition is reached. Therefore this model requires a change in chemical composition as well. With $\epsilon/2k = 0.07$ and $0.10^\circ\text{C}/\text{km}^2$, the crust contains so little radioactivity that the temperature does not flatten out and intersects the Clapeyron curve until greater depths are reached. For the pyrope transition, with $c = 1000^\circ\text{C}$, none of the models calculated provides a phase transition for depth in the range of 50 to 500 km. The only kinds of models that could have this phase transition in this range would be ones in which there was a very high concentration of heat sources in the first 1 or 2 km and almost no radioactivity below this depth.

For the albite transition, with $c = 50^\circ\text{C}$, the phase changes in the 30- to 50-km range occur for the model B_1 (with $\epsilon/2k = 0.321^\circ\text{C}/\text{km}^2$) at somewhat greater depths than for the pyrope transition. For the two-layer model B_2 , with the very high value of $\epsilon/2k$ in the upper crust, the transition occurs at a rather shallow depth. With $c = 250^\circ\text{C}$ none of the models indicates transitions at the depth of M ; again such

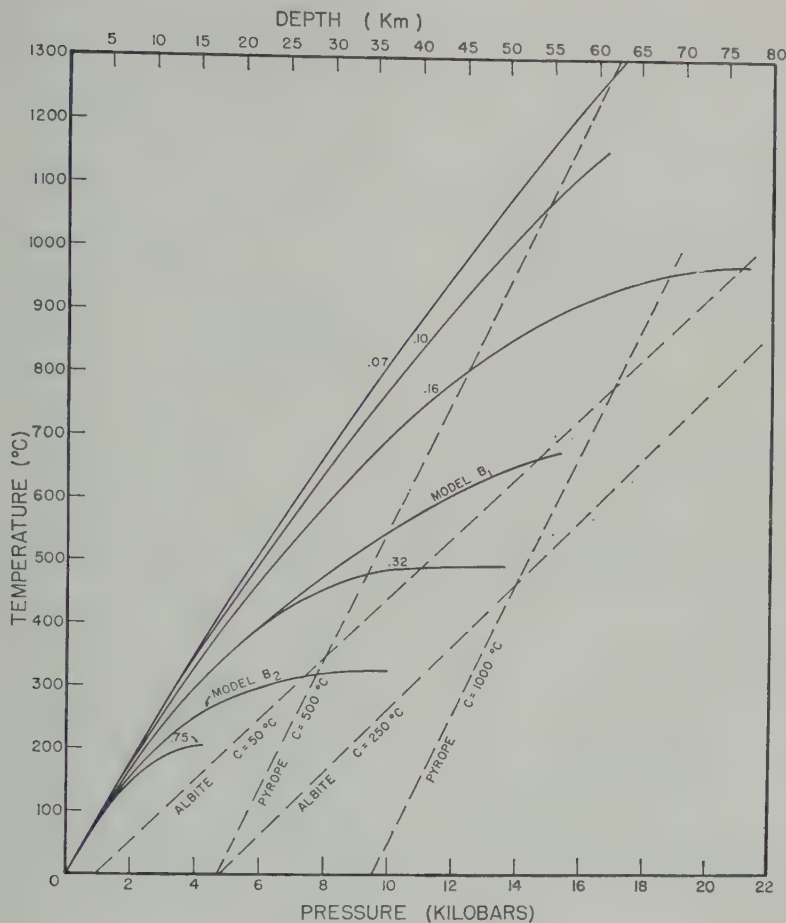


Fig. 1. Temperature-pressure curves for single- and double-layer models, the crustal thickness being undefined by the model. The intersection of one of these curves and one of the dashed Clapeyron curves for a given phase transition gives the pressure and temperature at the Mohorovicic discontinuity appropriate to this model and phase transition.

tions could be obtained only by an unreasonable concentration of all the sources of the heat flow within the first 1 or 2 km.

oceanic crust. While these calculations are consistent with the hypothesis that phase transitions which have been observed in the history take place at the depth of the continental M, it is more difficult to make this same transition with regard to the oceanic crust, as has been discussed by previous authors [Robertson and Birch, and MacDonald, 1957; Birch and Albee, 1960; MacDonald and Ness, 1960]. Calculations have been made for model C, oceanic crust, shown in Figure 3. The results of this calculation, again using equation 3, are

shown in Figure 4, labeled 'normal oceanic crust,' together with the curve previously discussed for the two-layer continental crust. The temperature-depth curve does not intersect any of the pyrope or albite phase-transition curves. Even if one gives up the attempt to use these measured Clapeyron curves and makes up Clapeyron curves to give intersections at the required depths such as those marked 'phase change A' and 'phase change B,' it is difficult to obtain reasonable thicknesses for both continental crust (model B) and oceanic crust using the same phase change. For example, phase change A was drawn so as to intersect the temperature curve for the continental model at a depth of

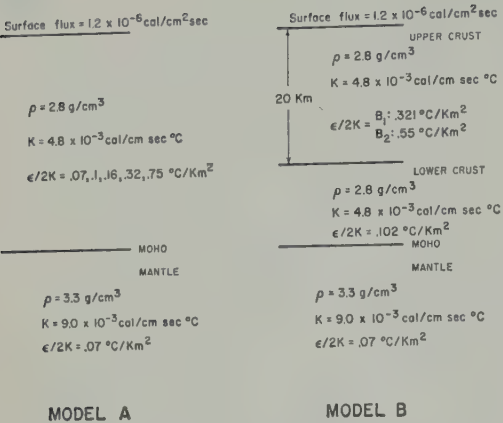


Fig. 2. Single- and double-layer crustal models used in the calculations.

35 km, a reasonable thickness for a continental crust. Its intersection with the temperature curve for the oceanic model is at the surface and would give no oceanic crust at all. If the intersection of phase change A with the model B curve is held fixed but the slope of the Clapeyron curve decreased, it can be made to intersect the curve representing oceanic temperatures. However the intersection is such that the low-density modification occurs at greater depths and the high-density modification at the surface. Phase change B, constructed to give a reasonable depth of 5 km for the oceanic crust, intersects the continental curve at a depth of only 8.4 km, which, of course, is much too shallow. By holding T and P constant at the base of the oceanic crust and decreasing the slope of phase change B, the situation can be somewhat improved. It is interesting to see how much improvement is possible. The most that the slope can be decreased while maintaining the high-density modification below the low-density modification is where the Clapeyron curve goes through the point (0°C, 0.5 kbar) as well as the base of the oceanic crust at (150°C, 20 kbar). This then determines the parameters in the equation $T = b\rho - c$ to be $b = 100^\circ\text{C/kb}$ and $c = 50^\circ\text{C}$. If the base of the continental crust is taken to be at 9.6 kb, the temperature at its base must be 910°C. Such high temperatures at this depth correspond (Fig. 1) to crustal models having a radioactivity slightly less than that of basalt, even in the upper portion of the crust.

These difficulties can also be removed by

modifying the oceanic model so that the properties of the oceanic crust are significantly different from conventional ones. For example, the conductivity of the oceanic crust can be assumed to be double the value used in model C. For temperatures above 100°C this value is greater than that of any material listed in the *Handbook of Physical Constants* [Birch, 1942] except quartzite, although a few basic and ultrabasic rocks approach this value. As is shown in Figure 5, the *ad hoc* 'phase change A' now gives a reasonable oceanic crustal thickness of 5.7 km and a continental crustal thickness of 35 km. Another way to obtain a similar result is to assume that the oceanic crust is much more radioactive than granite. In this case (Fig. 6) the thickness of the oceanic crust is found to be 7.3 km by using phase change A. It may be concluded that the same phase transition can produce both the oceanic and continental models if the phase transition has very special properties and somewhat unreasonable crustal models are used.

Effects of increased heat flow. One of the attractive features of a phase-transition M, pointed out by Kennedy, is that it provides a mechanism for causing significant changes in crustal thickness and, therefore (by isostasy) changes of elevation without the necessity of horizontal movement of large masses of crustal material.

This effect will be illustrated by the calculation of steady-state temperatures in an oceanic

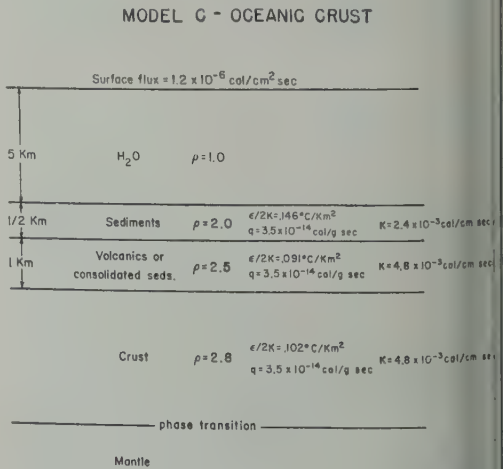


Fig. 3. Model of the oceanic crust used in the calculations.

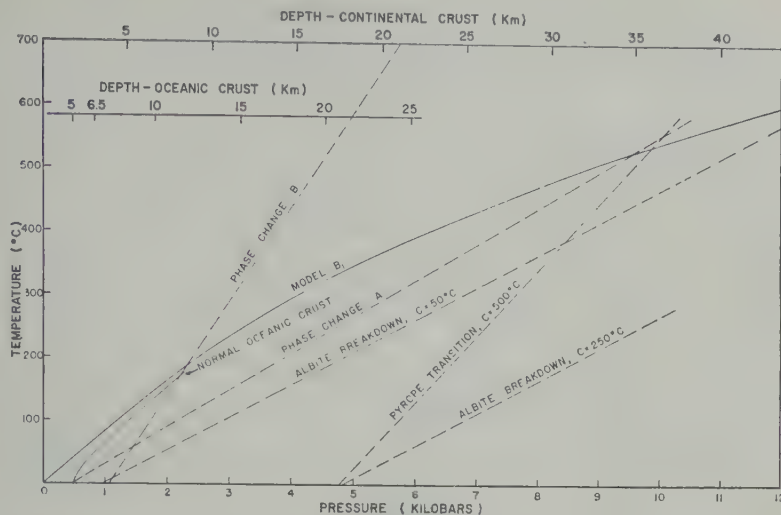


Fig. 4. Temperature-pressure curves of oceanic and continental crustal models together with Clapeyron curves of several phase transitions.

t, assuming that this crust is bounded by a phase transition. The results of calculating equilibrium temperature distribution for an oceanic crust with 1.5 times the average heat flow are shown in Figure 7. The water layer was omitted from the model; the reason for this omission will be made clear in the result. Again, in the previous section, neither the albite nor pyrope Clapeyron curves intersect the temperature-depth curve. If the *ad hoc* 'phase

change B' is used, the oceanic crust is found to have a thickness of 28 km. With the 1.5 km of consolidated and unconsolidated sediments assumed in the model, the total thickness will be 29.5 km. Such a crust will have its surface at sea level when in isostatic equilibrium with the normal oceanic crust (model C). Thus with phase change B an increase of 50 per cent in the heat flow suffices to transform an oceanic crust into a crust of nearly continental thick-

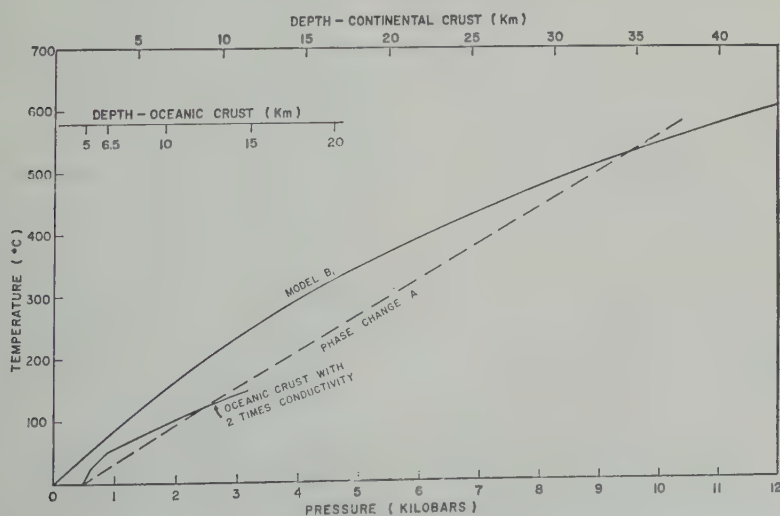


Fig. 5. Temperature-pressure curve for an oceanic crust with twice the thermal conductivity assumed in the standard model.

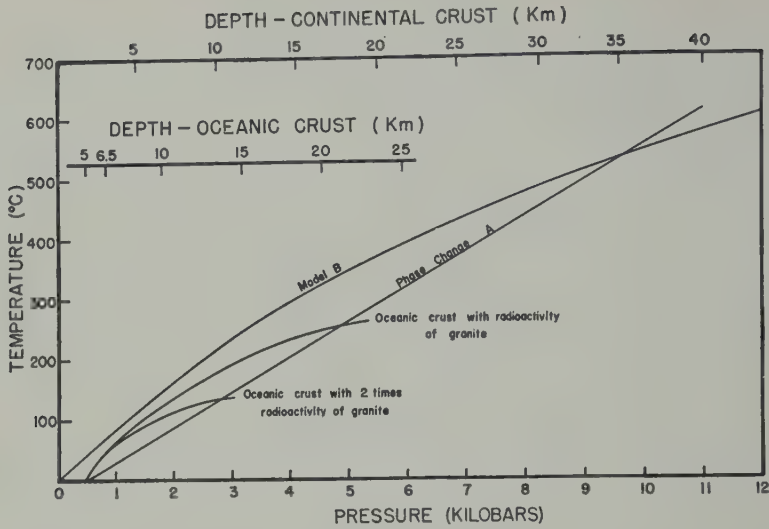


Fig. 6. Temperature-pressure curves for oceanic crusts containing more radioactivity than that assumed for the standard model.

ness. A similar calculation for a continental model (B_1) yields a 2-km increase in elevation for a 25 per cent increase in heat flow.

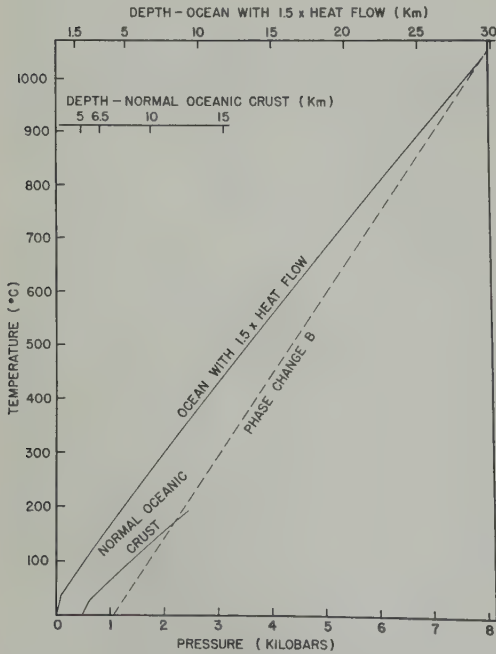


Fig. 7. Oceanic crust with surface heat flow increased by a factor of 1.5, the model otherwise the same as Figure 3.

Von Herzen [1959] reported heat flow values as high as 8 times the average value over a long belt in the southeast Pacific Ocean. The elevation of the sea floor is only about 1 km in this region, and no crustal thickening is detected by seismic methods. Even if the crust in this region is far from thermal equilibrium it is difficult to understand why there is no crustal thickening if the oceanic M is actually a phase change similar to phase change B. The problem could be alleviated somewhat by imposing even further *ad hoc* conditions on the phase transition, namely that its latent heat be extremely small, resulting in a much steeper Clapeyron slope. However it is necessary to carry the approach to an extreme in order to reconcile the phase-change hypothesis with Von Herzen measurements. These problems do not arise if the sources of the heat are high in the crust, e.g. a recent magma body within the sedimentary layer or a very unusual concentration of radioactivity near the surface.

Effects of differences in the concentration of radioactive nuclides. Comparison of analyses of rocks from ancient (2600 m.y. old) terranes and of younger rocks does not indicate any significant differences in the concentration of the radioactive elements potassium, uranium, and thorium. It would be expected, therefore, that the concentration of the radioactive nuclides

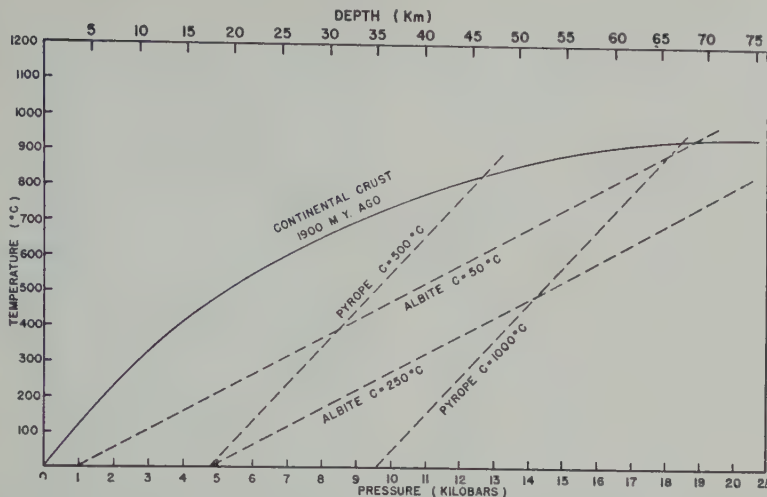


Fig. 8. Temperature-pressure curve for the continental crust 1900 million years ago, showing the effect of the increased radio-activity at that time.

now than it was in the past, owing simply to radioactive decay.

In particular, if the estimates of Birch [1955] for the relative contribution of these elements to the heat production in rocks of intermediate composition are used, it follows that 1900 m.y. ago the radioactive heat production in the crust was 1.42 times the present value. If at the present time the crust is bounded by a phase transition at a depth of about 35 km, in the past the higher crustal temperatures resulting from the higher rate of heat production would have caused the phase transition to occur at greater depths. The calculation of this effect is shown in Figure 8 for the double-layer model B. For the purpose of calculation it has been assumed that at the present time 0.216×10^{-6} cal/cm² sec of the surface heat flow is of nonradiogenic origin, and the remainder is produced by radioactive elements in the crust and beneath the crust. In calculating the temperature-depth curve for 1900 m.y. ago it is also assumed that all radioactive heat sources were 1.5 times as great, resulting in a surface heat flux of 1.70×10^{-6} cal/cm² sec, i.e. 1.42 times the present value.

Because of this increased radioactivity it may be seen from Figure 8 that the phase transition boundaries are displaced to much greater depths.

At the present time a continental block corresponding to model B₁, with the crust bounded by the albite transition at 48.2 km, is in isostatic

equilibrium with an oceanic crust corresponding to model C, then the continental block will now have an elevation of 1.1 km with respect to sea level. As discussed in the previous section, it is difficult to understand the oceanic M as a phase transition. If it is therefore assumed that the continental M is a phase transition and the oceanic M is not, then 1900 m.y. ago the continental block must have been 5.7 km above sea level. Similar, but less pronounced, effects result for the pyrope transition. This great increase in average continental elevation should have given rise to greatly different erosion rates in Precambrian times.

If, in spite of the difficulties discussed in the preceding sections, it is assumed that the oceanic M is a phase transition, the situation is not changed very much. Suppose that the oceanic crust is bounded by phase change B, which was contrived in order to provide the observed thickness of the oceanic crust at the present time. Then, making the same assumptions regarding the increased radioactive heat production 1900 m.y. ago, we calculate the temperatures in the oceanic crust and find that the phase transition occurs at a crustal thickness of 13.0 km. Making the isostatic calculation, we find that the average elevation of the continental crust above sea level is 4.5 km instead of the 5.7 km that was calculated for the case in which the oceanic M was not assumed to be a phase

transition. These effects will be less for phase transitions with Clapeyron curves of greater slope, such as the pyrope transition, and these phase changes will also be correspondingly less effective in producing changes in crustal elevation due to any other cause.

The formation of geosynclines and mountain ranges. We finally turn to one of the most interesting of the suggestions that have been made with regard to the properties of a phase-transition M, namely the effect of changes in the position of the crustal boundary on the formation of geosynclines and their subsequent elevation to form mountain ranges [Kennedy, 1959]. Of course the actual processes involved must be much more complex, since this mechanism does not, for example, provide an explanation for characteristic features of an orogenic belt such as intense deformation. However, rough calculations can be made to illustrate the role a phase change would play in this more

complex process. In brief, the cycle starts with the deposition of sediments in a water-filled trough. The weight of the sediments acts as a piston on the material below it, causing the temperature-depth curve to shift to the right. The intercept of this curve with the Clapeyron curve of the phase transition albite, ($c = 50^\circ\text{C}$) then takes place at a higher pressure, thinning the crust. Although retarded by the latent heat of the phase transition, the thinning may in some cases be sufficiently great to cause the trough to become deeper as the sediments are deposited. Eventually, however, the heating effects of the increased radioactivity in the crustal column together with any blanketing effects due to the lower conductivity of unconsolidated sediments will cause the temperature at depth to rise sufficiently to halt the thinning of the crust. If the temperature continues to rise, the crust will thicken and the bottom of the trough will start to rise even more rapidly than it is being filled

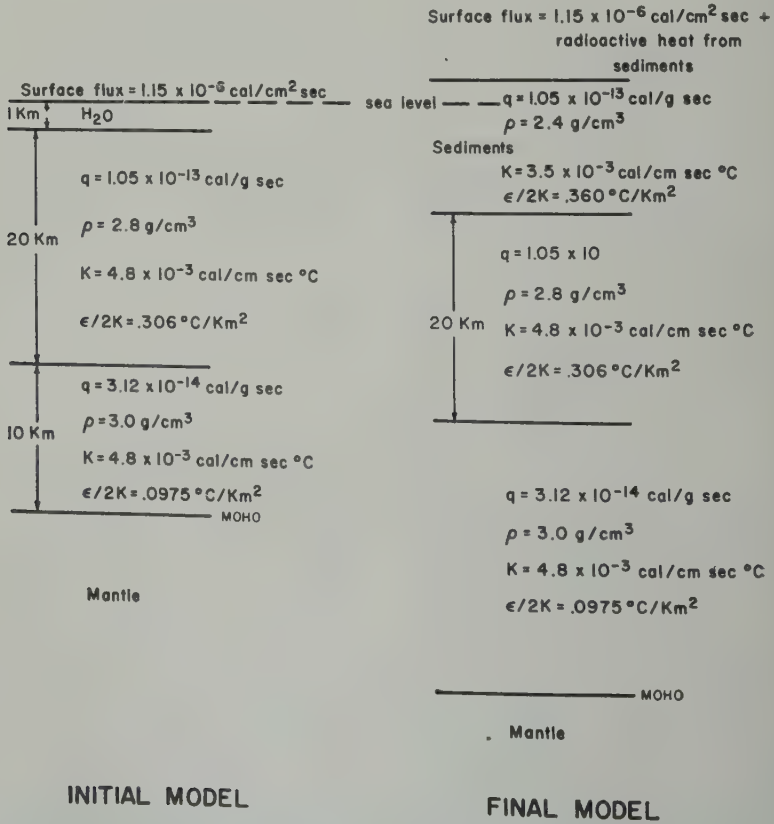


Fig. 9. Models used for the calculations of the temperatures and elevations produced by the deposition of sediments in a geosyncline.

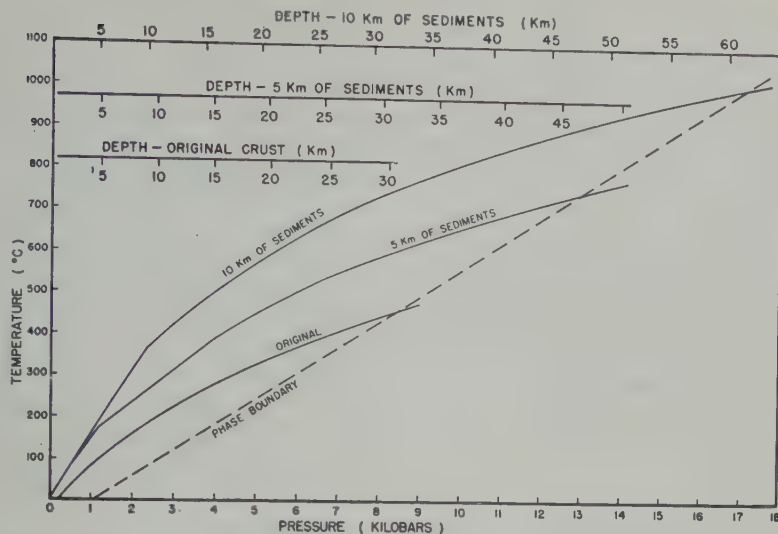


Fig. 10. Temperature-pressure curves showing the equilibrium temperatures following the deposition of 5 and 10 km of sediments.

sediments. As the temperature at depth is still higher, the surface will rise above sea level and sedimentation will cease. Eventually the temperature at depth will approach a new equilibrium at higher temperatures, and the crust will therefore be thicker than the original. In isostatic equilibrium the surface may be well above sea level.

This problem has already been treated at length by MacDonald and Ness, who obtained time-dependent solutions by use of an IBM 704 computer. There are important aspects of this problem to which the steady-state calculations as obtained in the present paper provide no information; the initial deepening of the trough and its subsequent filling are in an essential way time-dependent phenomena. They must be treated in all their complexity by numerical computations as has been done by MacDonald and Ness. However, as recognized by MacDonald and Ness, the difficulties of this approach forced these authors to make simplifying assumptions, such as neglecting the effects of isostasy and neglecting the heat carried by mass transport. The necessity for these simplifying assumptions limits the utility of the results.

Although, as mentioned above, the steady-state approximation is inadequate for a comprehensive discussion of this problem, it nevertheless

provides considerable insight into some aspects of it. This is done by calculating the isostatic elevations of two steady-state models: an initial state with a water-filled trough and a final state representing this trough filled with a given thickness of sediments in isostatic equilibrium with a normal oceanic crust. For a trough of a given depth there will be a correspondence between the final elevation and the thickness of the sedimentary layer. The maximum elevation of a mountain produced by the proposed mechanism is obtained, since the actual height will always be less because of erosion. Therefore, neglecting erosion, a given height will require the deposition of a calculable thickness of sediments.

The mechanism is illustrated by the models shown in Figure 9. The initial model is a two-layer crust with a total thickness of 30 km. The density of the lower crust is taken to be somewhat greater than that of the upper crust. When in isostatic equilibrium with the oceanic crust (model C), the surface of this crust will have an elevation of -1.0 km referred to sea level. Accordingly, it forms a trough filled with 1 km of water. In the model of the final state, the thermal conductivity of the sediments has been assumed to be slightly less than that of the deeper rocks, in order to introduce some thermal blanketing caused by only partially consolidated sediments. A hypothetical phase transition (phase

transition C) is chosen to intersect the temperature-depth curve (Fig. 10) at a depth of -31 km. This phase transition is similar in its general features to the albite ($\epsilon = 50^\circ\text{C}$) curve used previously, but it has a slightly greater slope. The final equilibrium states, after deposition of 5.0 and 10.0 km of sediments in the trough, are also shown in Figure 10. In these final states the phase transition takes place at 46.7 km, below the surface if there are 5 km of sediments and at 61.7 km for 10 km of sediments, and the final elevation will be +1.7 km for 5 km of sediments and +4.0 km for 10 km of sediments. Thus, for sedimentary thickness found in geosynclines, quite appreciable elevations may be expected.

The principal question left unanswered by these equilibrium calculations is the thickness of sediments that can be deposited before the trough fills. The equilibrium calculations give no information on this matter. The order of magnitude of the time required for heat to diffuse through the 30 km of crust is 3×10^7 years. By this time the phase transition will be migrating downward, thickening the crust, and filling the trough from the bottom. The numerical calculations of MacDonald and Ness, neglecting isostasy, indicate that the trough will fill in about 10^7 years at an assumed rate of sedimentation. According to Kay [1955], rates of sedimentation in geosynclines seldom exceed 500 meters in 1 million years, so it would be unusual if more than 5 km of sediments were deposited in 10^7 years. However the extent of the initial deepening of the trough and the time available before the trough is filled is sufficiently uncertain that the possibility of 10 km of sediments being deposited cannot definitely be excluded. It is also necessary to consider the rate of erosion of the emergent surface. The rate of uplift will be about 1 meter in 10^4 years if the duration is 4×10^7 years. The estimate of 0.3 meters per 10^4 years for the rate of erosion of the drainage basin of the Missouri River [Gilluly, 1952] indicates that significant erosion will take place during the uplift. If the erosion rates differed by an order of magnitude from these estimates, the surface might have been eroded as fast as it was uplifted, or on the other hand it might be uplifted without severe erosion. To make further progress in solving this problem it appears necessary to obtain time-dependent solutions in

which the effects of isostasy are considered.

APPENDIX

The origin of the differences between the results of this paper and those of MacDonald and Ness are briefly discussed.

All statements made by these authors regarding the relationship between the temperature at a given depth and the fraction of the earth's heat originating in the crust are incorrect. The exact way in which this error originated in their work is not clear, as there are several ambiguities in the definitions of the quantities entering into their expressions.

In calculations based on isostasy the effect of the deposition of the sediments in water was not considered. Algebraic errors also lead to the incorrect conclusion that the thickness of the sediments that can be deposited is independent of the initial depression.

The preceding objections do not apply to the second part of the paper, in which time-dependent solutions are obtained. However, it may be pointed out that neglecting the convective term in the heat flow equation (due to the conversion of crustal material to mantle material, and vice versa, as the phase boundary moves) may be significant. The convective heat flow will be given approximately by

$$\begin{aligned} V \cdot \nabla T &\sim V_{\text{Moho}} \frac{(\rho_M - \rho_c)}{\rho_c} \left(\frac{\partial T}{\partial x} \right)_{\text{Moho}} \\ &\sim \frac{10 \text{ km}}{10^7 \text{ yr}} \left(\frac{0.5}{2.8} \right) \left(\frac{10^\circ}{\text{km}} \right) = 1.7 \times 10^{-6} \text{ }^\circ\text{C/sec} \\ &= 6 \times 10^{-14} \text{ }^\circ\text{C/sec} \end{aligned}$$

This may be compared with the source term ϵ/ρ_c , which for basalt is about $1.3 \times 10^{-13} \text{ }^\circ\text{C/sec}$. Thus the convective term is of the same order of magnitude as the source term.

Acknowledgments. I wish to thank G. J. Wetherill, D. T. Griggs, and G. C. Kennedy for their critical discussion of this material.

REFERENCES

- Birch, Francis G., *Handbook of Physical Constants*, Geol. Soc. Am. Spec. Paper 36, 1942.
- Birch, Francis G., *Physics of the Crust*, Geol. Soc. Am. Spec. Paper 62, p. 114, 1955.
- Birch, F., and P. LeComte, *Temperature-pressure*

- ne for albite composition, *Am. J. Sci.*, 258, 217, 1960.
- y, J., A. C. Waters, and A. O. Woodford, *Principles of Geology*, W. H. Freeman and Co., Francisco, p. 135, 1952.
- Marshall, *Sediments and Subsidence through*, *Geol. Soc. Am. Spec. Paper 62*, p. 665, 1955.
- edy, G. C., The origin of continents, mountain ranges, and ocean basins, *Am. Scientist*, 47, 91, 1959.
- ing, J. F., The nature of the Mohorovicic discontinuity, *Trans. Am. Geophys. Union*, 39, 955, 1958.
- Donald, G. J. F., and N. F. Ness, Stability of phase transitions within the earth, *J. Geophys. Research*, 65, 2173-2190, 1960.
- Robertson, E. C., F. Birch, and G. J. F. MacDonald, Experimental determination of jadeite stability relations to 25,000 bars, *Am. J. Sci.*, 255, 115-137, 1957.
- Thompson, J. B., Jr., The thermodynamic basis for the mineral facies concept, *Am. J. Sci.*, 253, 65, 1955.
- Von Herzen, R., Heat-flow values from the southeastern Pacific, *Nature*, 183, 882, 1959.
- (Manuscript received December 12, 1960; revised June 20, 1961.)



The Use of Hornblendes and Pyroxenes for K-Ar Dating¹

STANLEY R. HART

*Department of Terrestrial Magnetism, Carnegie Institution, Washington, D. C., and
Department of Geology and Geophysics, Massachusetts Institute of Technology,
Cambridge, Massachusetts*

Abstract. The K-Ar ages of 12 hornblendes, one actinolite, and two pyroxenes were determined. When these ages are compared with ages of associated biotite, feldspar, or zircon, good agreement is found in most cases. No evidence is found for the existence of 'excess' radiogenic argon in these hornblendes. A maximum limit of 5×10^{-7} cc STP/g can be placed on possible 'excess' radiogenic argon in one sample. The potassium content of the amphiboles and pyroxenes is high enough so that Paleozoic or older samples can be easily dated, using present techniques. The rubidium content of the hornblendes is too low to be generally utilized for Rb-Sr dating.

INTRODUCTION

As have provided the basis for most K-Ar geochronologic mineral age determinations done since 1947. Their use has become accepted almost everywhere with the exclusion of other possible minerals. With the available techniques, it is feasible to measure K-Ar ages on minerals in which potassium occurs as a minor or even a trace element. It is probable that some of these low potassium minerals could prove more resistant to argon loss under metamorphic conditions than do micas.

A study of mineral ages across a contact metamorphic zone [Hart, 1961] showed that the hornblendes there retained argon much better than the biotite. Amirkhanov, Bartnitskii, and Voitkevich [1959] measured the diffusion coefficient of argon in a pyroxene and found it to be very small at geologic temperatures. This evidence seemed to warrant further study of amphiboles and pyroxenes.

The alkali cation position in hornblende is partly filled, and Damon and Kulp [1958] suggested that this might contain original 'excess' radiogenic argon, as for the beryl, tourmaline, and cordierite that they studied. To investigate this possibility, and to compare the argon retentivity of hornblendes and pyroxenes with micas, a number of samples from reasonably well-dated localities were analyzed for potassium and argon.

SAMPLE PREPARATION

In most instances mafic mineral concentrates were available as starting materials. Various size fractions from -60 to +200 were used. These were purified on a magnetic separator and in adjusted methylene iodide. Biotite was the dominant impurity. Grain counts were made on all samples. In addition, X-ray diffraction patterns were run on all samples and compared with patterns made from known mixtures of hornblende and biotite. The detection limits for biotite using the X-ray method was about 1 per cent. In several cases, the X-ray examination revealed intergrown biotite impurity where little free biotite was detected in the grain counts. This impurity was partly removed in one instance by dealing with a -270 +325 mesh-size fraction. More extensive purification procedures were not attempted, and in all cases the biotite impurity contributed less than 40 per cent of the total potassium in the sample. Portions for K and Ar analysis were taken by mixing and sampling at random with a spatula.

The location, rock type and estimated purity of the samples are given in Table 1.

EXPERIMENTAL TECHNIQUES

Argon analysis. An alundum crucible containing 3 to 5 grams of sample was placed in a Kanthal-wound resistance furnace. This was enclosed with a water-jacketed bell jar. The samples were heated in vacuum at 300° to 350°C

TABLE 1. Location and Description of Samples

Sample Number	Locality	Rock Type	Estimated Biotite Impurity, %	Contribution of Potassium from Biotite Impurity, %
HB47*	Wheaton, Md.	Tonalite	2	35
HB21*	Ellicott City, Md.	Granite	8	30
HP20*	Devault, Pa.	Baltimore gneiss	0.5	3
HSK*	Bear Mtn., N. Y.	Storm King granite	<0.2	< 1
HG15*	Crossnore, N. C.	Granite gneiss	2	25
HB15†	Shenandoah, Virginia	Gneiss	<0.2	< 1
H3089	Dill Sta., Ontario	Gneiss	<0.2	< 1
P3073	Kenogami, Quebec	Gabbro	<0.1	< 5
H3070	Chicoutimi, Quebec	Syenite	6	30
A3426	Neelon Twp., Ontario	Sudbury gabbro	<0.1	< 8
H3006	Rockport, Mass.	Cape Ann granite	5	20
H3136	Oak Bay, N. B.	Granite	<0.2	< 3
P3069	Mont Royal, Quebec	Tinguaite dike	<0.1	< 7
H3451	Dryden, Ontario	Amphibolite	5	40
H4068	Eldora, Colorado	Idaho Springs schist	<0.2	< 1

* Hornblende concentrates furnished by G. R. Tilton and G. L. Davis, Geophysical Laboratory, Washington, D. C.

† Hornblende concentrate furnished by B. R. Doe, Geophysical Laboratory, Washington, D. C.

All other samples obtained from the MIT Age Studies project.

H, hornblende; A, actinolite; P, pyroxene (3069-augite, 3073-hypersthene).

for about an hour. Tests of this procedure showed negligible radiogenic argon loss from a hornblende when heated for 30 minutes at 425°C, and less than 2 per cent loss when heated for 4 hours at 625°C. After this outgassing, the samples were fused directly by heating at 1100° to 1200° for 30 minutes. Enriched argon 38, calibrated against the argon content of air, was added, and the released gases were purified using hot copper oxide and titanium sponge.

The isotope ratios were measured using a 4½-inch 60° Reynolds-type mass spectrometer. The standard errors of 6 to 8 scans for the 40/38 and 36/38 ratios were 0.1 and 0.3 per cent, respectively. A correction for the residual at mass 36 was always less than 10 per cent. Argon 36 from the spike accounted for about half the total argon 36 peak. Tests of the 36 residual with hydrogen indicated that no enhancement took place under normal conditions. The discrimination of the ion source was frequently checked, using purified air argon. Air argon blanks in the extraction system were normally in the range 1 to 2 × 10⁻⁵ cc STP and accounted for most of the air argon observed in the purified argon samples.

One of the samples (P3069) was also run in

a system that allowed dropping 1 to 2 gram samples onto outgassed NaOH flux in a nickel furnace. The air-argon level of samples run in this system ranges from 1 to 5 × 10⁻⁵ cc STP.

The reproducibility of argon analyses was in the range 1 to 2 per cent. The analyzed argon content of the MIT standard biotite (3.85 × 10⁻⁴ cc STP/g) agrees within 1 per cent with that reported by other investigators [Hurley, Fairbairn, Pinson, and Faure, 1960].

Potassium-rubidium analysis. The potassium concentrations were determined in each case by both flame photometry and isotope dilution. The flame photometer solutions were prepared by decomposing 0.5 to 1.0 gram samples in sulfuric and hydrofluoric acids and diluting to 1 liter. As an internal standard 800 ppm lithium was added, and the solutions were compared with standards on a Perkin Elmer model-146 flame photometer, using propane flame. The unknowns and standards were both made up in 0.045 N H₂SO₄. Blanks for the procedure yielded about 25 micrograms of potassium. Corrections made for this were always less than 2 per cent. Six duplicate solutions indicated a precision of about 2½ per cent for the flame photometer potassium analyses.

TABLE 2. Analytical Data for Amphiboles and Pyroxenes

Sample number	Potassium, wt. %		Rubidium, ppm	Sodium, wt. %	Ar ⁴⁰ * (in 10 ⁻⁵ cc STP/g)	Ar ⁴⁰ * Total
	Flame Phot.	Isotope Dil.				
347	0.394, 0.403	0.397	83.5	0.92	0.543	0.29
321	2.12	2.05	68.8	...	2.65	0.71
20	1.35	1.34, 1.29	16.3, 15.6	0.92	6.65	0.83
SK	1.66, 1.69	1.68	36.1	1.25	7.82	0.82
115	0.532, 0.542	0.508, 0.531	12.1, 13.4	0.54	2.20	0.62
115	...	1.66	31.0	1.15	7.66	0.91
089	1.31	1.29	15.1	...	6.31	0.87
073	0.155	0.148, 0.148	0.56, 0.53	0.24	1.04	0.50
070	1.05, 1.15	1.33, 1.33	56.9, 61.0	...	4.58	0.80
426	0.0914	0.0872	2.05	...	0.925	0.26
006	1.49	1.68, 1.68	82.2, 79.8	1.14	2.26	0.56
136	0.482	0.506	10.3	...	0.740	0.28
069	0.112	0.102	4.40	...	0.0693, 0.0694	0.36, 0.31
451	0.663	0.835, 0.824	54.5, 51.7	0.65	15.7	0.92
068	1.01, 1.06	1.00	9.84	0.96	6.56	0.85

* indicates radiogenic component.

The isotope dilution procedure involved decomposing 0.5 gram samples in perchloric and hydrofluoric acids and spiking suitable aliquots with a mixed K⁴¹ and Rb⁸⁷ spike. A portion of the aliquot was applied to the single filament of a 6-inch solid-source mass spectrometer and the ⁴⁰K/⁴¹K and Rb⁸⁵/Rb⁸⁷ ratios were determined in a single run. Each ratio is the mean of 10 scans. The standard error of the potassium ratios ranged from 0.05 to 0.25 per cent, and of the rubidium ratios, from 0.1 to 0.4 per

cent. The ³⁹K/⁴¹K ratio was measured at low current before the rubidium ratios, and later at high current after the rubidium ratios. The low current ratios averaged about 1 per cent higher than the high current ratios. The reason for this is not known but it could arise either from initial contamination or sample fractionation. In some cases, the rubidium ratios were determined only by using an electron multiplier. Duplicate runs from the same aliquot generally produced the isotope ratios to within 1 per

cent. As a further test of accuracy, the alkalis from a number of these same aliquots were separated by sodium tetraphenyl boron precipitation. The isotope ratios of these separated alkalis also agreed within 1 per cent with the ratios obtained from the unseparated portions. Six com-

plete duplicate isotope dilution analyses indicate a standard deviation of 1.7 per cent for the potassium analyses and 4.5 per cent for the rubidium analyses.

Contamination blanks for the isotope dilution procedure showed from 3 to 6 micrograms of potassium and from 5×10^{-3} to 5×10^{-4} microgram of rubidium. The sodium tetraphenyl boron step added less than 0.1 microgram of potassium and 0.01 microgram of rubidium.

Sodium analysis. After taking the aliquot for the K + Rb analysis, the remaining solution was diluted so that the sodium concentrations were in the range 1 to 10 micrograms/ml. These were compared with standards on a Beckman model DU spectrophotometer with flame attachment. Duplicate samples indicate the standard deviation to be about 5 per cent. No attempt was made to evaluate the accuracy of the sodium determinations.

Constants used.

$$\lambda K^{40} = 5.30 \times 10^{-10} \text{ yr}^{-1}$$

$$\lambda eK^{40} = 0.585 \times 10^{-10} \text{ yr}^{-1}$$

$$K^{40} = 1.22 \times 10^{-4} \text{ g/gK}$$

Other mineral ages recomputed to these constants, and

$$\lambda Rb^{87} = 1.39 \times 10^{-11} \text{ yr}^{-1}$$

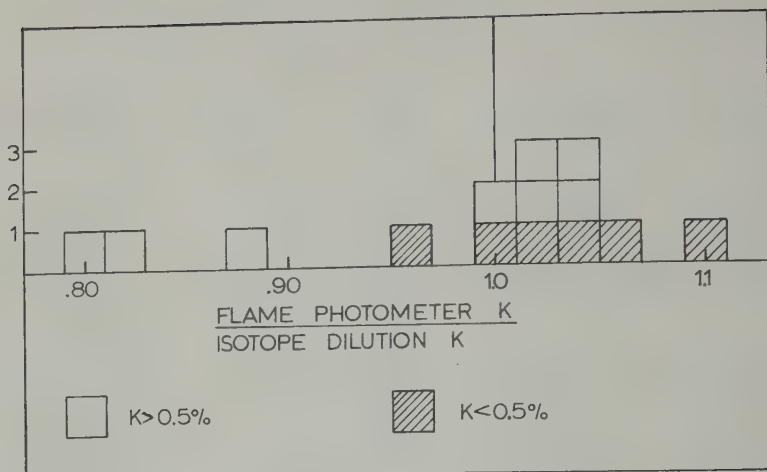


Fig. 1. Agreement of flame photometer and isotope dilution analysis of potassium in amphiboles and pyroxenes.

RESULTS AND DISCUSSION

The analytical data are presented in Table 2. Figure 1 shows a comparison of the flame photometer and isotope dilution potassium results. It is apparent that variations exist which are well outside the precision of the individual measurements. Flame photometer analyses for potassium in micas have been shown to be quite reliable [Hurley, and others, 1958]. However, interference from other elements may be more serious in hornblendes where, for example, the Ca/K ratio is commonly 200 times larger than in micas. Without further studies of interference and absorption phenomena the flame photometer analyses for potassium in these hornblendes and pyroxenes are considered less reliable than the isotope dilution analyses. The isotope dilution results are believed to be accurate within 1 or 2 per cent and are used in the age calculations.

It is probable that much of the rubidium in these hornblendes is contributed by the biotite impurity. Because of the general low level of rubidium in the pure hornblendes, they will generally be unsuitable for Rb-Sr dating. Pinson, Fairbairn, and Cormier [1958] did find hornblendes that contained as little as 10 ppm normal strontium. These might yield meaningful Rb-Sr ages if sufficiently old.

The results of sodium analyses (Table 2) on nine of the samples show that the atom ratio, Na/K, is greater than unity in each case. The

sum of sodium plus potassium represents the extent to which the alkali cation position is filled. In terms of an ideal tremolite structure, the alkali cation positions for these nine hornblendes range from 30 to 80 per cent filled. This is the range usually observed for natural hornblendes [Boyd, 1959] and indicates that these samples do have structural holes where 'excess' radiogenic argon could be held.

The calculated K-Ar ages are presented in Tables 3 and 4, along with published ages for coexisting minerals. In comparing the hornblende ages with ages on these coexisting minerals, two points should be considered. First, are any of the hornblende ages anomalously high, suggesting 'excess' radiogenic argon? And second, what resistance to argon loss during metamorphic events do the hornblende ages imply relative to the other minerals?

The first two samples in Table 3 are thought to be Paleozoic intrusives from the Maryland Piedmont, where the last metamorphic event affecting biotite ages seems to have occurred about 300 m.y. ago [Tilton, Wetherill, Davis, and Hopson, 1958; Tilton, Davis, Wetherill, Aldrich, and Jäger, 1959]. The hornblendes were apparently also affected by this metamorphism, giving K-Ar ages equal to K-Ar and Rb-Sr ages on the biotite. Because of the relatively low potassium and low age of the Wheaton, Maryland, hornblende, its age would be noticed

TABLE 3. Hornblende K-Ar Ages Compared with Ages of Associated Minerals

Sample number	Locality	Zircon						
		Hornblende K-Ar	Biotite		238	235	207	232
			K-Ar	Rb-Sr	206	207	206	208
1	Granite Ellicott City, Md. (a)	300 ± 10	315	290	355	370	450	310
7	Tonalite Wheaton, Md. (b)	315 ± 15	495	510	570 ± 50	485
0	Baltimore gneiss Devault, Pa. (c)	960 ± 30	1010	485
	Storm King granite Bear Mtn., N. Y. (c)	910 ± 25	840	930	960	990	1060	850
5	Granite gneiss Crossnore, N. C. (a)	840 ± 25	690	720	800 ± 50	680
5	Gneiss Shenandoah, Va. (c)	900 ± 25	800	880	1070	1100	1150	1110

ferences for ages other than hornblende K-Ar: (a) Tilton, Davis, Wetherill, Aldrich, and Jäger, 1959; Davis, Tilton, Aldrich, Wetherill, and Bass, 1960; (c) Tilton, Wetherill, Davis, and Bass, 1960.

affected by 5×10^{-7} cc STP/g of 'excess'ogenic argon. This may be considered a maximum limit for possible excess argon in this particular sample.

The last four samples of Table 3 are from the Alleghenian orogenic belt. A number of Precambrian rocks from this belt have been shown consistently give mineral ages in the 900–1100 m.y. range [Tilton, Wetherill, Davis, and Bass, 1960]. Allowing for some decrease in apparent age because of later, possibly Paleozoic, metamorphism, these four hornblendes are in general agreement with this range.

The biotite from Devault, Pennsylvania, is one of several known instances where the K-Ar age is significantly higher than the Rb-Sr age; also Sudbury gabbro, Table 4. The comparable K-Ar age in both cases is the same as the biotite K-Ar age. An explanation of the discordant biotite age pattern based on incorporation of radiogenic argon at the time of the metamorphism therefore does not seem likely. It is possible that a high external activity of radiogenic argon during metamorphism kept the argon diffusion rate low enough so that appreciable loss was prevented from both the hornblende and the biotite.

The Bear Mountain, New York, and Shenandoah, Virginia, samples show mica and hornblende ages that are slightly lower than the concordant zircon ages. Both hornblende ages are higher than the corresponding biotite

K-Ar ages and both are the same as the biotite Rb-Sr ages.

There are no biotite ages for the Crossnore, North Carolina sample. However, a number of biotites from western North Carolina have shown Rb-Sr and K-Ar ages in the range 300–400 m.y. [Long, Kulp, and Eckemann, 1959]. The Crossnore zircon ages are discordant but consistent with an original age in the range 900–1100 m.y. The hornblende age is somewhat lower than this presumed original age, but it is significantly higher than biotite ages in the near vicinity.

The geologic ages of the following localities are less well-documented, but still provide useful comparisons. The first three samples in Table 4 are from the Grenville Province of eastern Canada. Mineral ages from this province invariably fall in the range 900–1200 m.y. The hornblende age of the Dill Station, Ontario, sample is in reasonable agreement with the biotite ages and within the usual range of Grenville ages. The sample from Kenogami, Quebec, has no coexisting mineral ages. The pyroxene K-Ar age from this sample, however, is within the usual range of Grenville ages. The biotite and feldspar Rb-Sr ages on the Chicoutimi, Quebec, sample are typically Grenville, but the hornblende age of 700 m.y. is clearly low. No unique explanation for this can be given on the basis of the single sample.

The sample of Sudbury gabbro from Neelon

TABLE 4. Hornblende, Actinolite, and Pyroxene K-Ar Ages Compared with Ages of Associated Minerals

Sample Number	Locality	Hornblende	Biotite		Feldspar
		K-Ar	K-Ar	Rb-Sr	Rb-Sr
H3089	Gneiss	940 \pm 30	960	915	...
	Dill Sta., Ontario (a)				
P3073	Gabbro (pyroxene)	1240 \pm 75
	Kenogami, Quebec				
H3070	Syenite	710 \pm 30	...	950*	840
	Chicoutimi, Quebec (b)				
A3426	Sudbury Gabbro (actinolite)	1660 \pm 100	1780†	1330†	...
	Neelon Twp., Ontario (a)				
H3006	Cape Ann granite	310 \pm 10
	Rockport, Mass.				
H3136	Granite	335 \pm 15	300 \pm 60
	Oak Bay, New Brunswick (c)				
P3069	Tinguaita Dike (pyroxene)	163 \pm 10
	Mont Royal, Quebec				
H3451	Amphibolite	2370 \pm 80
	Dryden, Ontario				
H4068	Idaho Springs schist	1180 \pm 30	80	410	...
	Eldora, Colorado (d)				

* Reported previously as hornblende Rb/Sr Age.

† Biotite is from different sample of Sudbury gabbro. McKim Twp., Ont. (3196).

References for ages other than hornblende K-Ar: (a) Fairbairn, Hurley, and Pinson, in press; (b) Pinson, Fairbairn, and Cormier, 1958; (c) Hurley, 1958; (d) Hart, 1961.

Township, mentioned before, is another instance where the biotite K-Ar age is markedly older than the biotite Rb-Sr age. The actinolite K-Ar age is also high and, within experimental error, the same as the biotite K-Ar age.

The Cape Ann granite has been correlated with the Quincy granite, which is overlain unconformably by Pennsylvanian sediments. In the absence of other mineral ages, the hornblende K-Ar age of 300 m.y. can only be said to be reasonable.

The hornblende age from Oak Bay, New Brunswick, is in essential agreement with the feldspar Rb-Sr age. Other mineral ages are not available.

The Mont Royal, Quebec, sample is from a tinguaita dike on St. Joseph's Boulevard, Montreal East. The tinguaita is associated with the Monteregean intrusive there. No other mineral ages are available from Mont Royal, but several ages have been published for other of the Monteregean intrusives. A biotite K-Ar age of 122 m.y. was reported by the Canadian Geological Survey [Lowdon, 1960] for the Brome Mountain nordmarkite. A biotite from the essexite at Mt. Johnston gave a K-Ar age of 110

m.y. [Hurley, 1960]. The original age reported for the Mont Royal pyroxene of 100 m.y. [Hart, Fairbairn, Pinson, and Hurley, 1960] was in error because of the large atmospheric argon correction. The sample was rerun in duplicate using the sample dropping system described above. The results were in close agreement at 163 m.y. The difference between this age and the biotite ages from the other intrusives is in the direction of radiogenic argon contamination, either original 'excess' argon, or from impurity introduced during sample processing and purification. The excess is about 2×10^{-7} cc STP/g. The fact that 0.05 per cent of a 1000 m.y. biotite impurity could contribute this excess points to one of the major difficulties inherent in working with samples low in potassium. It is also possible, of course, that the Mont Royal intrusive is actually older than the other Monteregean intrusives.

The Dryden, Ontario, hornblende again has no ages of coexisting minerals for comparison. However, Aldrich and Wetherill [1960] have reported biotite ages from Sioux Lookout and Kenora, Ontario. Both localities are within 10 miles of Dryden. The ages (m.y.) reported for

tes from the two localities are, respectively, K-Ar, 2200 Rb-Sr, and 2490 K-Ar, 2550 Rb-Sr. There is evidence in the biotite ages for at least 2600 m.y. effect that has caused some of argon and strontium. The hornblende Rb-Sr age 2370 m.y. may also show evidence of such an effect.

The Eldora, Colorado, sample was collected from the Precambrian Idaho Springs formation and is about 54 m.y. from the contact of a 54 m.y. Laramide intrusion.

A series of ages in this contact zone has been reported previously [Hart, 1960]. The exponential argon retentivity shown by this hornblende was the first indication that hornblendes can be used to prove useful for K-Ar dating.

CONCLUSIONS

The determinations of (Na + K) demonstrate that these hornblendes have alkali-cation vacancies which are only partly filled. These natural holes provide sites where occluded or trapped radiogenic argon could be held. However, in no event does the K-Ar age of the hornblende or the actinolite exceed the probable thermal age of the host rock as deduced from the associated mica, feldspar, and zircon. There is no evidence from this study for the existence of 'excess' radiogenic argon in hornblende as postulated by Damon and Kulp [1960].

It is not clear from this study that hornblende will always retain argon better than biotite. In some cases, this is notably true; in others, the reverse seems to be true. The relationship of pyroxenes to micas is even less clear. In any event, the use of these minerals, and particularly hornblende, will permit dating of many intermediate and mafic rocks which have not previously been dated.

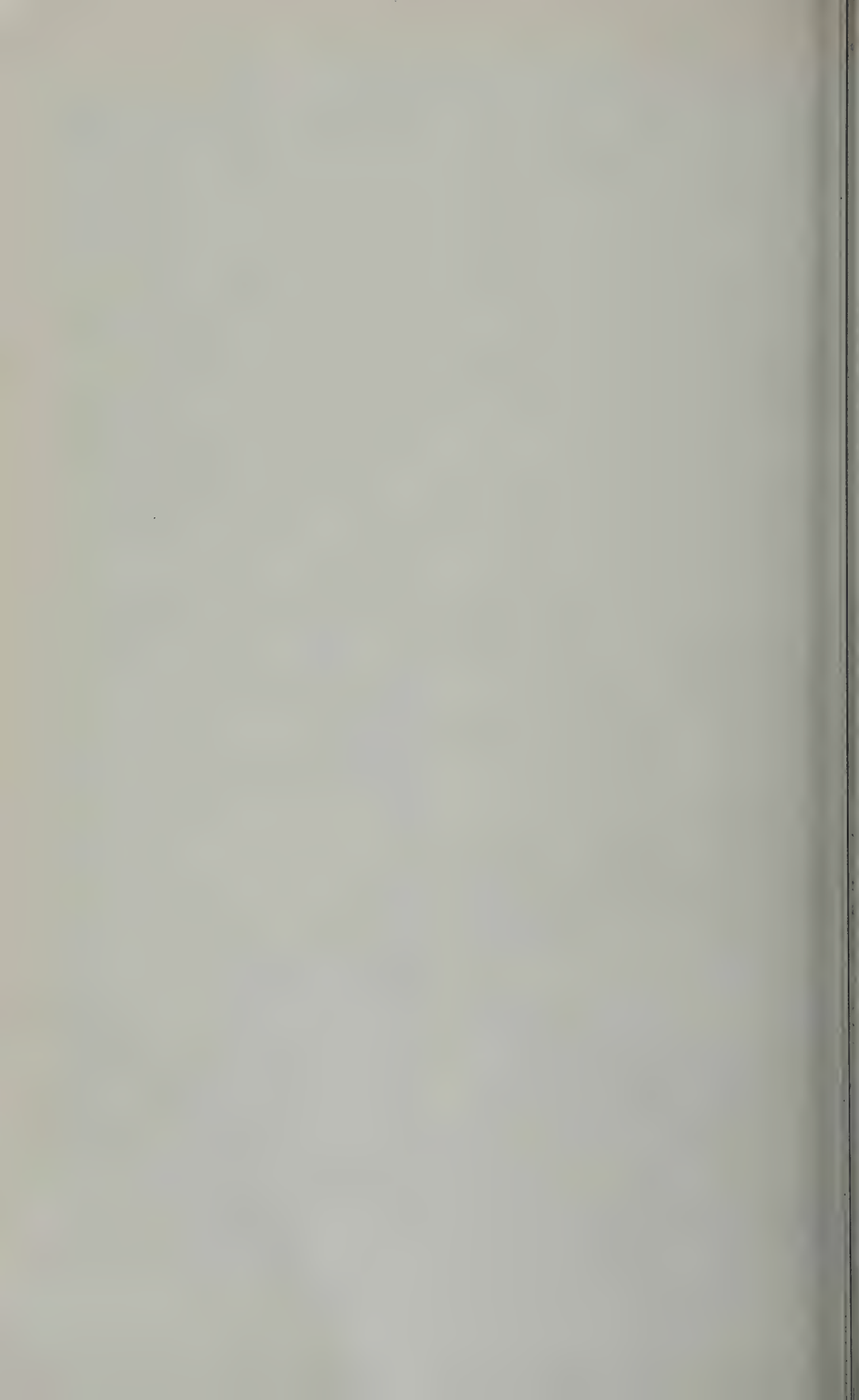
Acknowledgments. I wish to thank those who provided samples for this study. The interest and assistance of W. H. Pinson, H. W. Fairbairn, L. T. H. and G. L. Davis is gratefully acknowledged. The research was supported in part by contract AT(30-1)-1381. Finally, I am indebted to P. M. Hurley, whose encouragement and support made this work possible.

REFERENCES

Bartnitskii, S. B., L. T. H., and G. W. Wetherill, Rb-Sr and K-Ar ages of rocks in Ontario and northern Minnesota, *J. Geophys. Research*, **65**, 337-340, 1960.

- Amirkhanov, Kh. I., E. N. Bartnitskii, S. B. Brandt, and G. V. Voitkevich, On the migration of argon and helium in several rocks and minerals, *Doklady Akad. Nauk SSSR*, **126**, 160-162, geochemistry series, 1959.
- Boyd, F. R., Hydrothermal investigations of amphiboles, *Researches in Geochemistry*, Wiley and Sons, New York, 377-396, 1959.
- Damon, P. E., and J. L. Kulp, Excess helium and argon in beryl and other minerals, *Am. Mineralogist*, **43**, 433-459, 1958.
- Davis, G. L., G. R. Tilton, L. T. Aldrich, G. W. Wetherill, and M. N. Bass, The ages of rocks and minerals, *Carnegie Inst. Wash. Year Book*, **59**, 147-158, Washington, D. C., 1960.
- Fairbairn, H. W., P. M. Hurley, and W. H. Pinson, Mineral and rock ages at Sudbury-Blind River, Ontario, *Proc. Geol. Assoc. Can.*, 1961.
- Hart, S. R., Mineral ages and metamorphism, Conference on geochronology of rock systems, *Ann. N. Y. Acad. Sci.*, **91**, 192-197, 1961.
- Hart, S. R., H. W. Fairbairn, W. H. Pinson, and P. M. Hurley, Use of amphiboles and pyroxenes for K-Ar dating, Abstract, *Geol. Soc. Am. Bull.*, **71**, 1882, 1960.
- Hurley, P. M., and others, Variations in isotopic abundances of strontium, calcium, and argon and related topics, *MIT, 6th Ann. Prog. Rept. U. S. A. E. C. NYO-3939*, 1958.
- Hurley, P. M., H. W. Fairbairn, W. H. Pinson, and G. Faure, K-Ar and Rb-Sr Minimum Ages for the Pennsylvania Section in the Narragansett Basin, *Geochim. et Cosmochim. Acta*, **18**, 247-258, 1960.
- Long, L. E., J. L. Kulp, and F. D. Eckelmann, Chronology of major metamorphic events in the southeastern United States, *Am. J. Sci.*, **257**, 585-603, 1959.
- Lowdon, J. A., Age determinations by the Geological Survey of Canada, Rept. 1, Isotopic ages, Paper 60-17, *Dept. Mines and Tech. Surveys*, Canada, 1-51, 1960.
- Pinson, W. H., H. W. Fairbairn, and R. F. Cormier, Sr/Rb age measurements on hornblende and feldspar and the age of the syenite at Chicoutimi, Quebec, Canada, *Bull. Geol. Soc. Am.*, **69**, 599-601, 1958.
- Tilton, G. R., G. W. Wetherill, G. L. Davis, and C. A. Hopson, Ages of minerals from the Baltimore gneiss near Baltimore, Maryland, *Bull. Geol. Am.*, **69**, 1469-1474, 1958.
- Tilton, G. R., G. L. Davis, G. W. Wetherill, L. T. Aldrich, and Emilie Jäger, The ages of rocks and minerals, *Carnegie Inst. Wash. Year Book*, **58**, 170-178, Washington, D. C., 1959.
- Tilton, G. R., G. W. Wetherill, G. L. Davis, and M. N. Bass, 1000-million-year-old minerals from the eastern United States and Canada, *J. Geophys. Research*, **65**, 4173-4179, 1960.

(Manuscript received May 9, 1961; revised June 13, 1961.)



Natural Variations in Isotopic Abundances of Silicon

DAVID TILLES¹

*Department of Physics
University of California, Berkeley*

Abstract. Measurements of natural variations in isotopic abundance ratios of silicon are reported. A maximum natural range of 5.3 per mil in Si^{30} has been observed to date. Coexisting biotite, quartz, and feldspars from some specimens of igneous rock from the Yosemite region differ by as much as 3 per mil; in other grossly similar igneous rocks from the same region the same minerals differ by less than 0.3 per mil. In those rocks with large differences between the minerals, the ratio of $[\text{Si}^{30}]/[\text{Si}^{28} + \text{Si}^{29}]$ increases in the order biotite, quartz, feldspar. All granitic rocks studied are enriched in Si^{30} with respect to meteorites and basic rocks. Measurements on some sediments and biologic samples covered a range of approximately 4 per mil. A group of tektites covered a range of less than 0.7 per mil, with isotopic abundances centrally distributed within the observed normal terrestrial range. The results are interpreted, and applications of such studies to problems of geothermometry and formation of igneous rocks are discussed.

INTRODUCTION

Variations in isotopic abundance of the stable isotopes of silicon, the second most common element in the earth's crust, have received little detailed study. An unpublished set of measurements by Marsden is reported by Rankama [1954]. Two other groups of measurements, made independently by Allenby [1954] and Reynolds and Verhoogen [1953] using single ionization of ions, differed by about a factor of 2 in the observed difference between similar kinds of samples. Thus the heaviest sample from both of Allenby and of Reynolds and Verhoogen included marine samples of biogenic origin (diatomite, radiolarian earth, etc.). The latter were Yellowstone geyserite or sinter. Rankama, however, reported a difference between extreme samples approximately 5 times greater than that observed by Reynolds and Verhoogen.

The present study was begun with the hope of solving this discrepancy and with the expectation that greater precision and more detailed investigation might lead to useful applications of silicon isotope measurements. A preliminary abstract of some of the preliminary results has been presented previously [Tilles, 1959].

¹ Now at Smithsonian Astrophysical Observatory, Harvard College Observatory, Cambridge, Massachusetts.

EXPERIMENTAL APPARATUS AND PROCEDURES

Mass spectrometer. All measurements were made with a 60° mass spectrometer of 9-inch radius modified for rapid change from sample to standard gas and for null detection. The basic pattern of instrumentation was that used previously for measurements of carbon and oxygen isotopic abundance [McKinney, McCrea, Epstein, Allen, and Urey, 1950]. The material for mass spectroscopic analysis was SiF_4 gas.

The sample system consisted of two identical halves. Each half (Fig. 1) comprised a sample dumper, a sample decomposition tube, a decomposition furnace, a Toepler pump with large reservoir, a manometer, a capillary leak, and a bakeable metal valve connecting the line to the pumping manifold.

The entire system was of Pyrex, except for the valves, the capillary leaks, and the quartz tubing within which the BaSiF_6 was decomposed by heating. The sample decomposition tubes were individually surrounded by high-temperature cylindrical furnaces. Both the sample lines and the spectrometer source were arranged for convenient overnight bakeout to eliminate (1) spectrometer memory, (2) the possibility of exchange of silicon with that in the Pyrex and quartz of the spectrometer and sample line, and (3) trouble from background or impurity peaks. As an additional precaution, samples were baked out overnight in dumpers

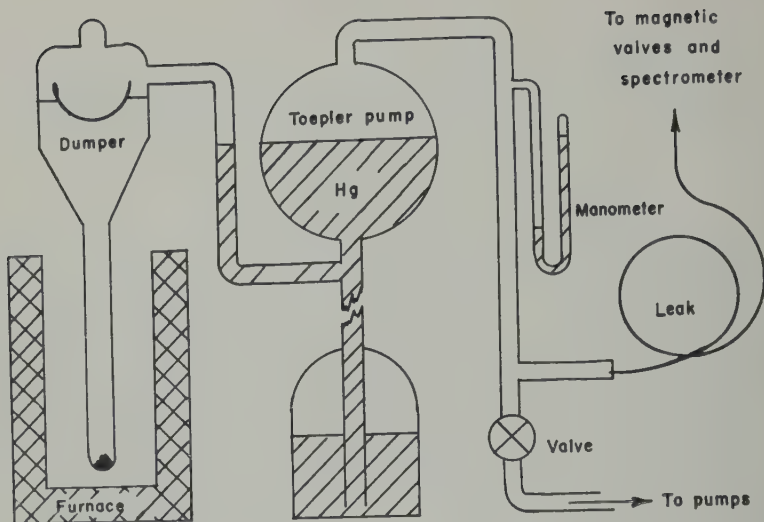


Fig. 1. Sample system (one of two identical sample lines).

so that the sample decomposition tubes could be baked out at higher temperatures.

The twin leaks and the magnetic valves were of conventional design. The electron bombardment gas source was also conventional, except that the final collimating slits were made narrow (.004 inch and .008 inch) and short ($\frac{1}{4}$ inch) in an attempt to get maximum possible resolution and to reduce beam scatter off the sides of the narrow portion of the analyzer tube where it passed through the magnet. The collector consisted of a slit plate to catch the mass 85 and 86 peaks while passing the 87 peak to a cup, with a defining entrance slit to block all but the three peaks of interest from the collector and two separate secondary electron-repeller plates adjacent to the slit plate and the cup.

The 87 peak was detected with a vibrating reed electrometer; the much larger 85 and 86 ion beams were detected with a d-c amplifier.

Chemistry and mineral separation. Silicon, extracted from samples by fusion with sodium carbonate flux, was converted into BaSiF_6 , which was decomposed by heating to give SiF_4 gas. This gas was admitted to the ion source of the spectrometer and the SiF_6^+ ions were used for mass spectrometric analysis. The essential procedures of the chemical extraction of silicon and its conversion to BaSiF_6 were similar to those described by Reynolds and Verhoogen [1953]. Whenever we started with known

amounts of silicon (standard preparation, p. quartz samples, etc.) and hence could precisely determine yield, we obtained yields between 80 and 100 per cent. Initial weights of mineral and final weights of BaSiF_6 were used to monitor yield.

BaSiF_6 was usually synthesized directly without intermediate separation, drying, and weighing of silica, and better over-all yields were obtained. As a check on the abbreviated chemical preparation, a few samples were prepared both ways—with separate extraction of silicon before preparation of BaSiF_6 and with direct preparation of BaSiF_6 from the initial sample. No detectable difference of δ was observed. (Definition of δ under Experimental Results.)

Unsatisfactory yields were obtained for a few of the more refractory samples studied. In an attempt to determine further the possible fractionation effects of incomplete attack of minerals, we performed a qualitative experiment in which fusion was purposely made incomplete so that only about 80 per cent of the silicon was removed from a dunite sample, then separated from the residue and carried through the normal high-yield chemical procedure. The observed shift in δ was 1 per mil heavier, leaving the residue depleted in Si^{80} . We left no material in our normal samples undissolved. With a few refractory samples, however, complete attack of the mineral required more lengthy procedures and resulted in some decrease of chemical yield.

ical effects of this kind may have been reliable for the considerable differences between a meteorite and olivine observed by *Reynolds and Verhoogen* [1953]. Although their results, based on input silica, were very good, it is possible that the procedures for converting more refractory minerals to silica may not have been quantitative. One cause of the much larger natural variations found by *Allenby* may have been laboratory fractionation in the proportion of SiF_4 during HF attack of the samples.

Experimental details. Additional checks and precautions, taken at different times during the course of measurements, included the following: (1) Duplicate runs with beam intensities $\frac{1}{4}$ of normal. (2) Duplicate runs with accelerating voltages differing by a factor 2. (3) Duplicate runs with sample reservoir pressures changed 10 percent from normal conditions. (4) Spectrometer memory checks with samples from Oak Ridge enriched silica, and having $\delta = 5.8$ per mil, run immediately after samples with $\delta = 0.5$ per mil. (5) Checking for mass fractionation with argon-nitrogen mixtures. (6) Checking of isotope exchange of samples with Pyrex or quartz parts of spectrometer, using enriched samples in sample lines for 24 hours. (7) Reheating of solid residue for 1 hour at 1000°C , after normal BaSiF_6 decomposition, to prevent additional evolution of SiF_4 or change in δ . (8) Checking of source memory and magnet valve leakage after every run. (9) Checking of accepted silicon isotope abundances by using a single collection. (10) Checking of secondary electron-repelling efficiency in collector. (11) Occasional runs with identical BaSiF_6 in sample and standard reservoirs. (12) Occasional runs with independently prepared samples from the same sample.

Within the sensitivity of the measurements (approximately 2 per cent for the experiments described above) no shifts in δ were observed. As a further precaution, both sample and standard BaSiF_6 were freshly decomposed from BaSiF_6 for each run, and in no case was SiF_4 stored for more than a few hours after preparation. Calibration samples were prepared by isotope dilution of Oak Ridge enriched Si^{30} with our primary standard. Measured values differed from the computed values by less than 0.2 per cent. As a further check both on our precision and

on that of the earlier measurements of *Reynolds and Verhoogen* we made new runs on two of the extreme samples studied by them. These samples, our numbers RV-15 and RV-17 (their numbers 15 and 17), were freshly prepared from the same natural materials and analyzed according to our normal methods. Each sample was run against our standard. The agreement with the earlier measurements attained by less precise methods was gratifying. The net difference between the two samples was 3.00 ± 0.06 per mil, as compared with the *Reynolds-Verhoogen* results of 3.1 ± 0.9 per mil. These probable errors represent for each difference the total of the probable errors for the two samples.

All probable errors shown in the following section include estimates of all sources of error (including incomplete yield and possible fractionation in the chemistry). The *Reynolds and Verhoogen* probable errors are statistical estimates based both on internal scatter during a single run and on the differences between two runs on the same sample. It is important, however, to note that their duplicate spectrometer runs were usually made from the same sample of BaSiF_6 ; thus their quotations of probable errors did not include an estimate of uncertainties from possible incomplete extraction of silica from the samples or incomplete yields.

Impurity peaks were found at masses 82, 83, and 84. These peaks, observed also by earlier investigators, are believed to be SiF_2O^+ , SiF_2HO^+ , and $\text{SiF}_2\text{H}_2\text{O}^+$, as suggested earlier by *Reynolds and Verhoogen*. If this is correct, one can compute the contributions at 85, 86, and 87 by isotopic molecules of the above, and with the observed magnitude of the impurity peaks the effect on the observed values of δ was never more than 0.01 per mil. If, as suggested by *Allenby*, the impurity represents the three silicon isotopes in SiF_2O , then the impurities are even less serious, as no contributions appear at masses 85, 86, and 87. The relative sizes of the three impurity peaks do not appear, however, to be consistent with such an assignment. As additional evidence of the correctness of the *Reynolds-Verhoogen* assignments, when one run was made in which the source was baked out at a much lower temperature than normal, the impurity peaks increased by an order of magnitude to about 0.1 per mil effect. Thus the

bakeout clearly helped eliminate impurity peaks. Possible hydride impurities would have been detected by a peak at mass 88, but no such peak was ever observed in the course either of this work or by Reynolds and Verhoogen.

More complete details of normal experimental procedures and precautions are given by *Tilles* [1960].

EXPERIMENTAL RESULTS

All measurements are given in values of δ , where δ is defined as

$$\delta_{\text{sample}} = \left\{ \frac{([Si^{30}]/[Si^{28} + Si^{29}])_{\text{sample}}}{([Si^{30}]/[Si^{28} + Si^{29}])_{\text{standard}}} - 1 \right\} \times 1000$$

All observations are presented relative to the standard silica used as our laboratory standard, a sample of vein quartz from the California Mother Lode region. For comparison with future work the δ value of standard sand (U. S. National Bureau of Standards Isotope Reference Sample Number 28, silica sand from the Oriskany deposit, provided by V. H. Dibeler) was measured on four occasions distributed throughout the period of this study. The four values, which were obtained with $BSiF_6$ from two independent chemical preparations, give an average value of δ for the NBS standard of $+0.29$ per mil with a probable error in the mean of 0.1 per mil.

The total natural variation of silicon isotope ratios observed was approximately 5 per mil. This extends the spread of about 3 per mil observed by Reynolds and Verhoogen but is much less than the 14 per mil found by Allenby.

Meteorites and basic rocks. Table 1 gives the results for two different dunites, a stony meteorite, and four Skaergaard samples [Wager and Deer, 1939]. Since yields on the Skaergaard samples were not satisfactory, the Skaergaard results are probably the least reliable of those reported. For comparison we have included in the last column of this table the earlier results of Reynolds and Verhoogen for different stony meteorite and a basalt. Their values are converted to our scale of δ values: matching our repeat runs on their extreme samples 15 and 17 (geyserite and diatomite) with their measurements and then simply taking differences from these extremes. (It should be pointed out that the definition of the Reynolds and Verhoogen δ values differs from ours. Their δ in terms of 30/28 ratios whereas our δ is measured directly as $30/(28 + 29)$. The correction is less than a few per cent of the value of δ , is not included here.) The 1 per mil difference observed between olivine and meteorite by Reynolds and Verhoogen [1953] may have been due to complications of a chemical nature. Since their measurements were made by comparison with SiF_4 from a single sample of $BSiF_6$, the differences between any pair of their measurements

TABLE 1. Dunites, Meteorites, and Basic Rocks of the Skaergaard Intrusion

Sample	Rock	Description	δ	δ (Reynolds and Verhoogen)
M-1	Meteorite	Richardton, N. Dak., chondrite	-0.32 ± 0.15	
13	Meteorite	Melrose, N. Mex., stone		-0.29 ± 0.10
0-1	Dunite	Jackson Cnty., N. Carolina (Reynolds Verhoogen standard)		
0-2	Dunite	Mt. Dunn, New Zealand	-0.30 ± 0.15	
7	Basalt	Augite from picrite basalt tuff, Mauna Kea, Hawaii	-0.35 ± 0.15	
SK-1	Fayalite	Skaergaard intrusion, Greenland		$+0.01 \pm 0.10$
SK-2	ferrogabbro	Greenland	-0.72 ± 0.4	
SK-3	Hortonolite	Skaergaard intrusion, Greenland	-0.09 ± 0.3	
SK-4	ferrogabbro	Skaergaard intrusion, Greenland	-0.53 ± 0.4	
	Olivine-free middle gabbro	Greenland		
	Hypersthene	Skaergaard intrusion, Greenland	-0.54 ± 0.3	
	olivine gabbro	Greenland		

will be unaffected by such problems, but their δ values with respect to olivine might be shifted by a constant amount.

Sierra Nevada whole rocks. Table 2 gives approximate weighted average silicon isotopic abundance of several igneous rocks from Sierra Nevada, based on available measurements on separated minerals. These rocks, in contrast to the dunites and meteorites, show significant differences in isotopic ratios. As a group, however, the more silicic rocks are clearly enriched in Si^{30} with respect to the silicic and ultrabasic rocks of Table 1. *Grant* [1960] predicted on theoretical grounds that if silicic and basic rocks originated by fractional crystallization of a common magma, the silicic rocks should be enriched in the heavy isotope of silicon with respect to the basic rocks and the magnitude of enrichment should depend upon temperature at which the differentiation occurred. A parallel increase in the fraction of the heavier isotope in the more silicic rocks was observed for oxygen in silicates by *Silverman* [1960] and has been more systematically studied by *Minogrodov, Dontsova, and Chupakhin*, [1960] and *Taylor and Epstein* [1960].

Tektites. Table 3 presents several measurements made on tektites. The isotopic composition is quite ordinary and within the terrestrial range. Only the bediasite is not within a narrow range between -0.3 and 0.0 per mil. The isotopic composition varies from the same as, to slightly lighter than, that of the meteorites and is similar to the one basalt measurement of *Reynolds and Verhoogen*.

A check on the possibility that tektites

TABLE 3. Tektites

Sample	Description*	δ
KA 352-2	Kalgoorlie,	$+0.03 \pm 0.1$
KA 352-2B	Western Australia	$+0.01 \pm 0.1$
KA 353-2B	Kalgoorlie,	
	Western Australia	-0.31 ± 0.1
KA 351-5	Kalgoorlie,	
	Western Australia	-0.17 ± 0.15
KA 358-3	Tan Hai, China	-0.20 ± 0.15
KA 359-3	Moldavite	-0.17 ± 0.15
KA 355-5	Bediasite, Lee	
	County, Texas	$+0.37 \pm 0.15$

* More complete descriptions have been published by *Reynolds* [1960].

originated from the impact of a comet on marine sediments [*Urey*, 1957] two samples of shale were measured. The silicon isotopic abundances of the shales, -0.23 and $+0.33$, are not inconsistent with the Urey theory. It is of interest that all the tektites have silicon abundance ratios similar to or slightly higher in Si^{30} than the meteorites. It has been suggested [*Cassidy*, 1956, 1958] that the tektites may have originated from breakup of an object having a chondritic interior and a tektitic crust. The observed difference of silicon isotope ratios between tektites and chondrites is parallel to the fractionation of these isotopes between silicic and basic crustal rocks of the earth. The present observations seem consistent with both the impact and the Cassidy hypothesis, but more measurements are needed.

Biologic and sedimentary samples. Observations on contemporary marine sponges and diatoms and some sediments appear in Table 4. We have included for comparison some related earlier measurements of *Reynolds and Verhoogen*, converted to our scale of δ values. These samples scatter very widely; the silicon in one of the sponge samples is the lightest observed. Biological processes are known to produce large isotopic fractionation in other elements and may also produce silicon isotopic enrichment.

Coexisting minerals in igneous rocks. Principal emphasis in this research was on study of coexisting minerals in igneous rocks. Results are summarized in Table 5 and Figures 2 and 3. The rocks are principally from Yosemite National Park and are described and mapped elsewhere [*Calkins*, 1930], and K-Ar ages are available for most of them [*Curtis, Evernden, and Lipson*,

TABLE 2. Sierra Nevada Whole-Rock Averages*

Rock	δ
Bay granodiorite	1.5 ± 0.2
Rock granite	1.2 ± 0.2
Sierran granite	0.7 ± 0.2
Central Peak granite	0.6 ± 0.2
El Capitan granodiorite	0.45 ± 0.3
Yosemite quartz monzonite	0.35 ± 0.2
Granite	-0.05 ± 0.2
Yosemite quartz monzonite	-0.2 ± 0.2

The whole-rock estimates are obtained from measurements on separated minerals together with known fraction of silicon in each mineral and ratios of mineral abundance in the respective whole-rock point counting [*Smith*, 1958].

TABLE 4. Biologic and Sedimentary Samples

Sample	Description	δ	δ (Reynolds-Verhoogen)
RV-17*	Marine diatomite, Miocene, Santa Barbara (rerun of Reynolds-Verhoogen sample 17)	1.66 ± 0.1	
2	Chert, partly recrystallized, Franciscan, California		1.66 ± 0
14	Chert, Monterey formation, Miocene, Berkeley, Calif.		1.16 ± 0
M-6	Diatoms, surface, 84°F water temp., Gulf of Panama	0.84 ± 0.2	
M-2	Shale, well near Marysville Buttes, California	0.33 ± 0.2	
M-1	Shale, Mt. Hamilton, Calif.	-0.23 ± 0.2	
12	Diatomite, fresh water, Calif.		-0.94 ± 0
16	Flint, from Dover chalk		-1.14 ± 0
M-5	Sponge, 200-280 fathoms, Baja, California	-1.27 ± 0.2	
RV-15*	Yellowstone geyserite (rerun of Reynolds-Verhoogen sample 15)	-1.34 ± 0.2	
M-4	Sponge, Gulf of California, less than 100-meter depth	-2.18 ± 0.2	

* Samples RV-15 and RV-17 are used to compare the Reynolds-Verhoogen values with those of this paper

1958]. The spread in isotopic ratios differs radically between the same sets of minerals from grossly similar rocks. The silicon ratios for several of the principal minerals were found to be identical, within a few tenths of a per mil, for rocks such as the Half Dome quartz monzonite and the Cathedral Peak granite. The corresponding differences in isotopic abundances were as large as 1 to 3 per mil in the Arch Rock granite, the Gateway granodiorite, and the El

Capitan granite. In all rocks with large differences the feldspars were always heavier than either the quartz or the mica. In four of the Yosemite rocks (the three mentioned above and the Hoffman pegmatite) the differences in silicon isotope ratios of quartz, feldspar, and biotite were significantly larger than uncertainties of measurement. In each of these rocks the isotopic ratios varied from lightest to heaviest in the order: biotite, quartz, feldspar.

Discussion. If complete equilibrium existed among all the minerals in a rock, the differences in observed values of δ between any pair of minerals for any particular element studied should be a monotonically decreasing function of equilibrium temperature [Urey, 1947; Bigeleisen and Mayer, 1947; Clayton and Epstein, 1959; Epstein, 1959]. True equilibrium of coexisting minerals is an uncommon phenomenon in natural rocks formed from melts. Usually some minerals will form earlier than others, fractionation of some minerals from the remaining melt will ensue, and cessation of isotopic change of different minerals with the melt will occur at different temperatures. Also, as crystallization proceeds, the isotopic composition of the uncrystallized fraction of the melt will shift in a complex way determined by the isotopic fractionation between each mineral and the melt and the amount of each mineral already crystallized

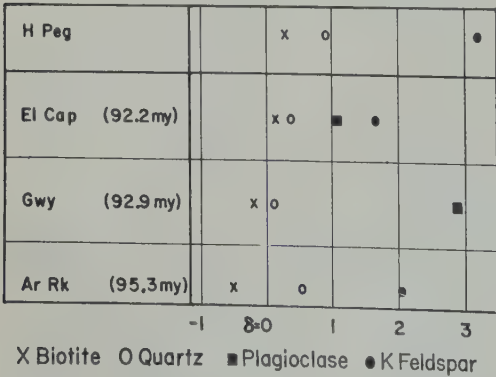


Fig. 2. Differences in silicon isotope ratios of separated minerals from igneous rocks. H Peg is Hoffman pegmatite, El Cap, El Capitan granite; Gwy, Gateway granodiorite; Ar Rk, Arch Rock granite. Numbers after rock name are K-Ar ages in millions of years determined from separated biotites. δ given in per mil.

temperature history of the mass of material, the temperature dependence of the equilibrium isotope fractionation between each mineral and the melt. Thus, some minerals may be chemically zoned just as some are chemically unzoned and this may be true even in the limiting case of isothermal crystallization if different minerals are not crystallizing simultaneously and at identical rates.

Isotopic differences that become 'frozen' in the minerals will reflect the temperatures at which exchange rates become negligible, and the temperatures will depend upon viscosity and reaction rates between the specific minerals in the melt. The isotopic differences will therefore indirectly reflect the effects of composition, water content, pressure, and other variables.

Reaction rates, however, typically display an exponential temperature dependence, so that a small difference in temperature may be responsible for a large difference in exchange rates. In some cases isotopic differences may therefore become 'frozen' over a moderate temperature range, and, if one studies the same minerals in different rocks, differences in the final temperature of isotopic exchange may be observable. Varying water content, pressure, and composition of starting materials it should be possible to grow synthetic mineral pairs—quartz and feldspar, for example—over a considerable range of temperatures and in this way calibrate experimentally both a silicon and an oxygen thermometer for crystallization temperature of silicate minerals in igneous rocks. Without either

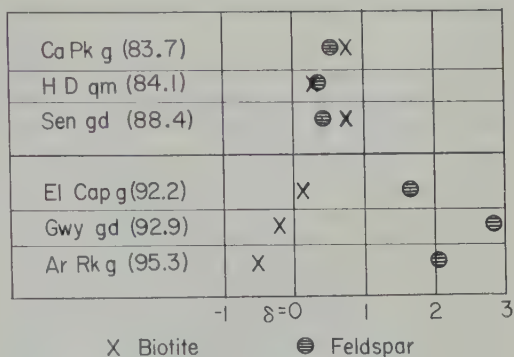


Fig. 3. Differences in silicon isotope ratios of biotite and feldspar from dated Yosemite intrusions. Ca Pk g is Cathedral Peak granite; H D qm, Half Dome quartz monzonite; Sen gd, Sentinel granodiorite; El Cap, El Capitan granite; Gwy gd, Gateway granodiorite; Ar Rk g, Arch Rock granite. Numbers after rock names are K-Ar ages in millions of years determined from separated biotites. δ given in per mil.

such empirical calibration or computations of equilibrium fractionation factors (extremely difficult for most of the complex silicates) any inferences about relative temperature of cessation of isotope exchange based on differences in observed isotope fractionation between mineral pairs are completely speculative.

Of the nine rocks for which data on both biotite and feldspar are available, only two (the Sentinel granodiorite and the Cathedral Peak granite) were found in which biotite was enriched in Si^{30} with respect to feldspar, and in neither of these two rocks were the differences

TABLE 5. Separated Minerals from Igneous Rocks of the Sierra Nevada

Rock	K-Ar Age, m.y.	δ Values for Separated Minerals			
		Biotite	Quartz	Plagioclase	K Feldspar
Half Dome quartz monzonite	143	-0.38 ± 0.15			-0.05 ± 0.15
Cathedral Peak granite	95.3	-0.51 ± 0.2	0.54 ± 0.2		2.05 ± 0.2
Gateway granodiorite	92.9	-0.20 ± 0.15	0.12 ± 0.2	2.86 ± 0.2	
El Capitan granite	92.2	0.12 ± 0.2	0.37 ± 0.2	1.05 ± 0.15	1.66 ± 0.2
Sentinel granodiorite		-0.16 ± 0.2			0.02 ± 0.15
Arch Rock granite	88.4	0.79 ± 0.2			0.42 ± 0.2
Half Dome quartz monzonite	84.1	0.31 ± 0.15			0.37 ± 0.15
Cathedral Peak granite	83.7	0.79 ± 0.15	0.68 ± 0.15	0.50 ± 0.15	0.55 ± 0.15
Gateway pegmatite	76.9	0.28 ± 0.2	0.92 ± 0.2		3.17 ± 0.3

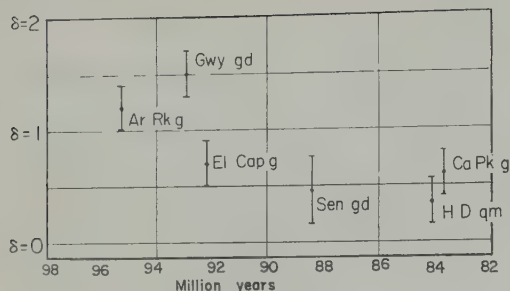


Fig. 4. Estimated whole-rock silicon isotope ratios of dated Yosemite plutons, based on available measurements on separated minerals. (For identification of symbols, see titles of Figs. 2 and 3). δ given in per mil.

much larger than the experimental uncertainties.

The observations reported here suggest that the normal pattern of silicon fractionation in igneous rocks may be biotite lighter than quartz, which in turn is lighter than feldspar. This pattern is distinctly different from the normal order of enrichment of the heavy isotope of oxygen in these same minerals. *Taylor and Epstein* [1960], studying separated minerals from the Southern California batholith, have consistently observed larger O^{18}/O^{16} ratios in quartz than in feldspar. There is, however, no intrinsic reason to expect fractionation of different elements in the same direction between the same pair of minerals.

The rather large differences in observed silicon isotope fractionation in coexisting minerals of different rocks (Table 5 and Fig. 2) are puzzling. Two possible explanations appear reasonable at this time:

1. The rocks with large silicon isotope differences were formed or continued isotope exchange at a much lower temperature than the rocks with small silicon isotope differences. This explanation requires not only considerable differences in temperatures of formation but also quite strong temperature dependence of silicon fractionation over the temperature range of interest. Determinations of the temperature dependence of isotope fractionation in silicates will show whether this explanation is reasonable. Further evidence on the possible differences in temperature of formation of these rocks can be obtained by oxygen isotope measurements on the separated minerals.

2. The rocks with large silicon isotope differences are extreme examples of disequilibrium assemblages of minerals, with final isotope frac-

tation determined by kinetic or other factors. Disequilibrium conditions and processes (for example, vapor transport mechanisms) might, under favorable circumstances, produce quite homogeneous isotope fractionation in a large rock mass. Parts of the Sierra batholith, like the El Capitan granite, display large vertical and horizontal exposures and have large endosilicon isotope differences between minerals that measure precisely. Such masses of rock are favorable for further study of this problem.

The mineral differences in the Arch Rock granite and the Gateway granodiorite (Table 5 and Fig. 2) are of interest: (1) The biotite in the Gateway is 0.3 per mil heavier than the biotite in the Arch Rock. (2) The quartz in the Gateway is 0.4 per mil lighter than the quartz in the Arch Rock. (3) The feldspar in the Gateway is 0.8 per mil heavier than the feldspar in the Arch Rock. (4) The whole-rock average for the Gateway is 0.3 per mil heavier than for the Arch Rock.

All the above results are consistent with the following model: (a) Both rocks originated from the same melt. (b) Both biotites crystallized from the respective melts at approximately the same temperature. (c) During the late stages of crystal growth, biotites ceased growing or exchanging isotopes with the melt at a significant rate before the quartz and feldspar ceased growing. (d) In the Gateway the quartz and feldspar continued isotope exchange down to a lower final temperature than in the Arch Rock, a result which might be expected if, for example, the Gateway had a greater water concentration than the Arch Rock.

It should be re-emphasized that the silicon and other stable isotope differences between the minerals are by themselves incapable of providing any information about the maximum temperature reached in the formation of a rock. They rather reflect the *lowest* temperature at which isotope exchange continued at a significant rate.

It is evident that the whole-rock average (Fig. 4) for the dated Yosemite plutons displays considerable scatter in the silicon ratios. Such bodies of igneous rock, if created by any process from pre-existing rocks, may reflect the isotopic composition of the sediments or other material from which they were formed and may have fairly wide scatter in isotopic ratios. It may be significant that the heaviest whole-rock values

The Yosemite series is that of the Gateway diorite, that the Arch Rock granite is also heavy, and that both of these are among the lightest of the rock masses which have been dated for the entire area. They are also on the extreme western edge of the entire igneous complex immediately adjacent to metamorphics and are the two oldest rocks of the Yosemite plutons which have been dated.

Large rock masses, if originating by complete melting, might be expected to smooth out local variations in 'inherited' isotope composition. Homogenization, however, might require isotopic transport over considerable distances, and its effectiveness would depend both upon the viscosity of the melt and the length of time it remained molten. The El Capitan granite, the Half Dome quartz monzonite, and the Cathedral Peak granite all have similar whole-rock values which would be the same within the limits of precision. A possible alternative explanation for the heavier whole-rock values of the Gateway diorite and the Arch Rock granite may be that the various plutons in the Yosemite region originated by differentiation from large masses of basic magma at depth. If silicic rocks formed by differentiation from a more basic magma, the silicic rocks should be heavier in ^{28}Si than the basic rocks [Grant, 1954]. A qualitative corollary of such a prediction is that the rocks formed at the higher temperature at which this separation took place should be the more enriched in ^{28}Si ; the result is that the lighter rocks should be. If the Gateway diorite and the Arch Rock granite were in equilibrium with a less polymerized silicate melt at depth at a lower temperature than the later rocks of the Yosemite series of plutons, the observed differences in whole-rock silicon ratios could be explained. Such an explanation is extremely speculative at this time. A lower temperature of differentiation for the oldest Yosemite granites is consistent with the possible lower temperature of final isotope homogenization suggested as one possible explanation for the mineral differences in these same rocks. Local melting of material with no opportunity for homogenization of isotope differences and the fractional differentiation mechanism are incompatible and cannot both be correct. The alternative interpretations of the isotopic evidence presented here, however, all suggest a high temperature of formation or final crystallization for the earlier Yosemite plutons than for the later ones.

Most of the rocks of the Yosemite region belong to two series of intrusions, studied and described by Calkins [1930]. The Half Dome quartz monzonite, the Cathedral Peak granite, and the Sentinel granodiorite are all included in what Calkins calls the Tuolumne intrusive series. The Taft granite (not dated by K-Ar methods and merely indicated by Calkins as younger than the El Capitan), the El Capitan, and the Arch Rock granite are all included in what Calkins refers to as the Yosemite Valley biotite granite series. The Gateway granodiorite is not clearly identified as part of either series, but from its age and location it would appear to belong with the Yosemite Valley series. The Guadalupe quartz monzonite is about 50 million years older than the Yosemite rocks and is considered to be unrelated to the main period of intrusion.

The three Yosemite rocks for which the oldest K-Ar ages were obtained all exhibit large silicon isotope differences between the minerals (Fig. 3). The three younger rocks studied exhibit small isotopic differences, and their average K-Ar age is about 9 million years younger than the older group. These two groups of rocks, corresponding to the separate series of rocks described by Calkins, thus display different patterns of silicon isotope fractionation, the reasons for which, as already mentioned, are not presently understood. These results, however, suggest the possibility of using isotope fractionation patterns as a supplement to petrologic and geochronologic methods of correlating related rock masses in different parts of a large batholith.

The Hoffman pegmatite, presumably formed from a residual fluid very rich in water or vapor, displays the large isotopic differences believed to be typical of either low temperature equilibration or disequilibrium effects. It is interesting to note the considerable similarity in isotopic patterns between this pegmatite and the early Yosemite rocks, despite their great dissimilarity in texture and crystal size. Another pegmatite was studied in greater detail, and the results are presented in a separate paper [Tilles, 1961].

SUMMARY

Measurements have been made of differences in isotopic abundance ratios of silicon from nat-

ural sources. The total observed natural range of approximately 5 per mil in $[\text{Si}^{30}]/[\text{Si}^{28} + \text{Si}^{30}]$ extends the limits of the earlier observations of Reynolds and Verhoogen but is still much less than the total variation reported by Allenby. Two dunities and a meteorite were found to have the same silicon isotope ratios. These and four rocks of the Skaergaard were all significantly lighter in Si^{30} than a group of granitic rocks of the Sierra Nevada. Measurements were made of silicon from several tektites. With the exception of one sample of bediasite from Texas, all were the same within limits of 0.3 per mil, and all were within the normal terrestrial range. A few measurements of sedimentary and biogenic silicon showed a wide scatter throughout the observed range of natural variation. The lightest sample found was silica from sponge spicules.

Differences in quartz, biotite, and feldspar silicon ratios of a group of rocks of the Sierra Nevada batholith were studied. In some of the rocks the abundance ratios are the same within a few tenths per mil; in others they differ by as much as 3 per mil. In all rocks where the differences between the minerals were large, the observed order of enrichment from lowest to highest Si^{30} abundance was: biotite, quartz, feldspar.

Present interpretation of these effects is extremely speculative. A more complete understanding of the measurements reported here will require additional data from both natural and laboratory systems. Present measurements do demonstrate, however, that differences in silicon isotope ratios can with care be measured with a precision of 0.1 to 0.2 per mil. They also strengthen suggestions of earlier studies that the natural variations of isotopic abundance ratios of silicon are large enough to provide a useful geologic tool.

Acknowledgments. I should like to thank J. H. Reynolds, who directed this research. G. Curtis and J. F. Evernden have been generous with both samples and encouragement. I am particularly grateful to Hugh Taylor for discussing his oxygen results and for providing many stimulating ideas. S. Silverman, C. McKinney, and H. Craig have been generous with technical details of high-precision mass spectroscopy. F. J. Turner, W. L. Quaide, E. D. Goldberg, G. V. Howard, and others have assisted by providing samples. A. Ryall and C. Perdue donated the quartz used as a laboratory standard. J. Lipson provided both encouragement and assistance during the early stages of this research. To O. Tobisch I owe thanks for his exten-

sive assistance in mineral separation and sample chemistry. Financial assistance was provided by the U. S. Atomic Energy Commission and the Institute of Geophysics of the University of California.

REFERENCES

- Allenby, R. J., Determination of the isotopic ratio of silicon in rocks, *Geochim. et Cosmochim. Acta*, **5**, 40-48, 1954.
- Bigeleisen, J., and M. G. Mayer, Calculation of equilibrium constants for isotopic exchange reactions, *J. Chem. Phys.*, **15**, 261-267, 1947.
- Calkins, F. C., The granitic rocks of the Yosemite region, in *Geologic History of the Yosemite Valley*, by F. E. Matthes, (*U. S. Geol. Surv. Prof. Paper 160*) 120-129, 1930.
- Cassidy, W. A., Australite investigations and the bearing on the tektite problem, *Meteoritics*, **426-437**, 1956.
- Cassidy, W. A., Achondrite investigations and the bearing on the origin of tektites, *Geochim. et Cosmochim. Acta*, **14**, 304-315, 1958.
- Clayton, R. N., and S. Epstein, The relation between $\text{O}^{18}/\text{O}^{16}$ ratios in coexisting quartz, calcite, and iron oxides from various geological deposits, *J. Geol.*, **66**, 352-373, 1958.
- Curtis, G. H., J. F. Evernden, and J. Lipson, The determination of some granitic rocks in California by the potassium argon method, *Calif. Dept. Nat. Resources, Div. Mines, Spec. Rpt.*, **54**, 1958.
- Epstein, S., Variations of the $\text{O}^{18}/\text{O}^{16}$ ratio in nature and some geologic implications, in *Researches in Geochemistry*, edited by P. H. Abelson, J. Wiley & Sons, New York, 1959.
- Grant, F. S., The geological significance of variations in the abundance of the isotopes of silicon in rocks, *Geochim. et Cosmochim. Acta*, **5**, 242, 1954.
- McKinney, C. R., J. M. McCrea, S. Epstein, H. Allen, and H. C. Urey, Improvements in mass spectrometers for the measurement of small differences in isotope abundance ratios, *Rev. Sci. Instr.*, **21**, 724-730, 1950.
- Rankama, K., *Isotope Geology*, McGraw-Hill Book Co., New York, 1954.
- Reynolds, J. H., and J. Verhoogen, Natural variations in the isotopic constitution of silicon, *Geochim. et Cosmochim. Acta*, **3**, 224-234, 1939.
- Reynolds, J. H., Rare gases in tektites, *Geochim. et Cosmochim. Acta*, **20**, 101, 1960.
- Silverman, S. R., The isotope geology of oxygen, *Geochim. et Cosmochim. Acta*, **2**, 26-42, 1958.
- Smith, A. R., The modal composition of the mafic intrusions of the Yosemite Valley, Unpublished Master's thesis, University of California, Berkeley, 1958.
- Taylor, H., and S. Epstein, $\text{O}^{18}/\text{O}^{16}$ ratios in coexisting minerals in igneous rocks (*Abstracts J. Geophys. Research*, **65**, 2528, 1960).
- Tilles, D., Natural variations in isotopic abundance

- s of silicon (Abstract), *Bull. Geol. Soc. Am.*, 688, 1959.
- D., Natural variations in isotopic abundances of silicon and their significance, unpublished Ph.D. thesis, University of California Physics Dept., Berkeley, 1960.
- D., Variations of silicon isotope ratios in a pegmatite, *J. Geophys. Research*, 66, 3015-1961.
- H. C., The thermodynamic properties of isotopic substances, *J. Chem. Soc.*, 562-581, 1947.
- H. C., Origin of tektites, *Nature*, 179, 556-1957.
- adov, A. P., E. I. Dontsova, and M. S. Chupakhin, The isotopic composition of oxygen in igneous rocks and meteorites, Transl. in *Geochemistry*, no. 3, 235-239, 1958 (Publ. in *Geokhimiya*).
- Vinogradov, A. P., E. I. Dontsova, and M. S. Chupakhin, Isotopic ratios of oxygen in meteorites and igneous rocks, *Geochim et Cosmochim. Acta*, 18, 278-293, 1960.
- Wager, L. R., and W. A. Deer, Geological investigations in East Greenland, 3, *Medd. Grønland*, 105, no. 4, 1939.

(Manuscript received October 19, 1960; revised June 13, 1961.)

Variations of Silicon Isotope Ratios in a Zoned Pegmatite

DAVID TILLES¹

*Department of Physics
University of California, Berkeley*

Abstract. Isotopic abundance variations of silicon in quartz and feldspar from the Rose Quartz pegmatite of the Pala district in Southern California are reported. All feldspar samples are enriched in Si^{30} with respect to adjacent quartz samples, and the magnitude of enrichment varies from 0.6 to 2.7 per mil in Si^{30} . Feldspar-quartz differences were found to increase from outer zones towards the inner quartz core. All measurements on quartz samples were within a total range of about 0.6 per mil. The feldspars most heavily enriched in Si^{30} are believed to result from Rayleigh fractionation during crystallization from a vapor phase.

Introduction. Measurements of natural variations of isotopic abundance of silicon have shown differences of 2 to 3 per mil may exist in coexisting minerals in a single hand specimen of fine-grained igneous rock [Tilles,

A small zoned pegmatite was selected for more detailed study, to investigate the following questions: (1) Magnitude of isotopic fractionation in a 'low temperature' igneous system. (2) Variation of magnitude of fractionation throughout the pegmatite body. (3) Mechanism of fractionation of a zoned pegmatite.

Examples were provided by Richard H. Jahns of a zoned pegmatite of the Pala district, California. This body, called the Rose Quartz pegmatite, is several hundred feet long and is surrounded by gabbro. A map of the pegmatite is shown in Figure 1. The sample locations are described more fully in appendix A. Quartz and feldspar were selected for detailed study because they were available together from all parts of the pegmatite. Locations of the quartz-feldspar pairs studied are shown in Figure 2.

Experimental procedures are discussed in detail elsewhere [Tilles, 1961] and were unchanged for these measurements, with one exception. The corrections reported here were made after a series of mishaps reduced the sensitivity of the low rates of the copper capillary leaks. A correction of -0.6 per mil was applied. Table 1 shows the runs made to determine the magnitude of the correction. After the correction,

work was done at Smithsonian Astrophysical Observatory, Harvard College Observatory, Cambridge 38,

completion of all the pegmatite measurements, runs were made in which the leaks were readjusted to give identical 87 peaks for radically different pressures on the upstream side of the two leaks. Nulls were then observed.

The results of these experiments suggested that the shift in all the nulls was actually produced by a small fraction of molecular flow back into the reservoirs. This effect, checked with a mixture of argon and nitrogen during early adjustment of the capillary leaks, was made negligible by the proper choice of leak dimensions, constrictions, and flow rates. It was not negligible, however, for the altered leak conditions during the pegmatite measurements.

It should be emphasized that, whereas this correction factor decreases somewhat the absolute reliability of the measurements reported here, the precision of the determination of differences between samples does not appear to be affected. These differences were of principal interest in the pegmatite investigation.

Experimental results. Table 2 and Figures 3, 4, and 5 give the measurements, all of which include the leak correction discussed above but are otherwise as measured in values of δ where δ is defined as

$$\delta_{\text{sample}} = \left\{ \frac{[(\text{Si}^{30})/(\text{Si}^{28} + \text{Si}^{29})]_{\text{sample}}}{[(\text{Si}^{30})/(\text{Si}^{28} + \text{Si}^{29})]_{\text{standard}}} - 1 \right\} \times 1000 \quad (1)$$

The laboratory standard was the same crushed vein quartz from the California mother lode used in the measurements previously reported [Tilles, 1961].

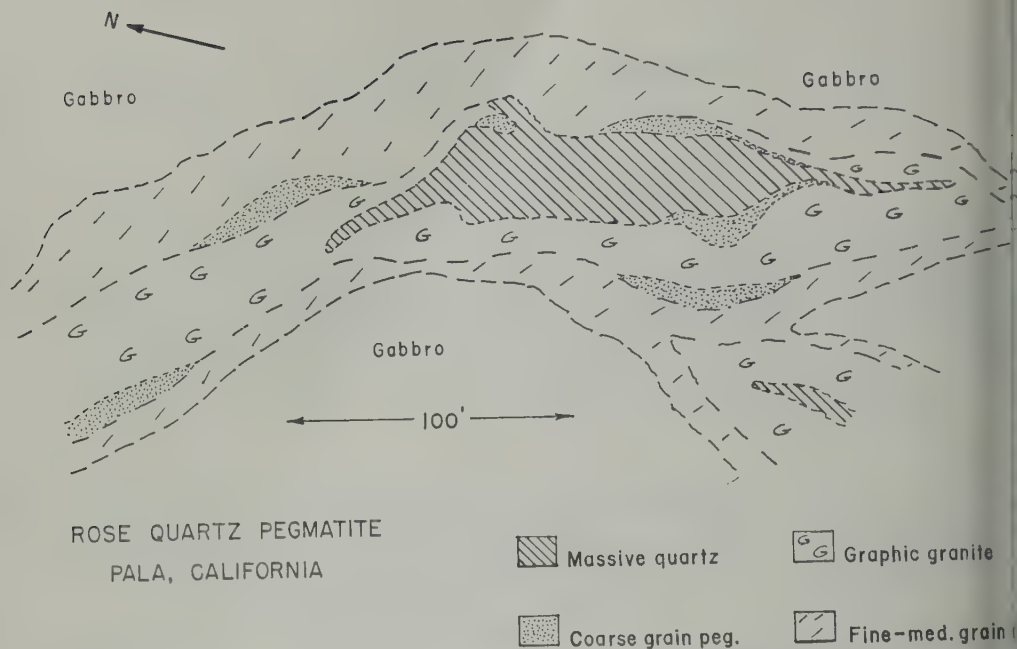


Fig. 1. Schematic map of Rose Quartz Pegmatite, Pala, California (based on original field mapping of R. H. Jahns).

Several facts are evident from these measurements:

1. In all cases the isotopic differences between adjacent quartz and feldspar increase from the outer edge of the pegmatite toward the inner quartz core. This is true both where the shifts are gradual over distances of 100 feet and where they change by 1 to 2 per mil over distances of only 15 or 20 feet.

2. The isotope values of the quartz samples from all sampled parts of the pegmatite are quite similar.

3. The big isotopic differences between adjacent quartz and feldspar are all a result of feldspar enrichment in Si^{30} and not of mutual

large shifts in both quartz and feldspar from some central value.

4. No clear correlation exists between magnitude of quartz-feldspar difference and crystal size. Thus, for example, sample RQ-7, a big grained, somewhat aplitic rock, displays a large quartz-feldspar difference smaller than any of the coarser grained graphic granite samples or even the coarse pegmatite samples RQ-1 and RQ-2 at the margin of the quartz core.

5. The largest quartz-feldspar difference observed in this pegmatite is no larger than the largest corresponding difference for the big grained Yosemite igneous rocks studied [Tilles, 1961]. The quartz-feldspar differences for all the other rocks in which they were measured are summarized in Table 3. As in these earlier measurements, feldspar was always found enriched in Si^{30} with respect to quartz when the differences were large enough to be significant.

6. The foot wall in the vicinity of the core determined by samples RQ-3 and RQ-8, shows markedly heavier values for the feldspar than does the hanging wall.

Possible models of pegmatite formation. From these results certain inferences about the foot

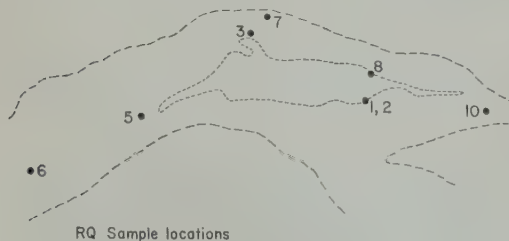


Fig. 2. Locations of samples.

TABLE 1. Leak Correction Data

Sample	Date Run Originally	Original Value of δ	Date Run for Correction	Value of δ	Correction to δ
J-2	4-29-59	+2.10 \pm .05	7-21-59	+2.75 \pm .05	-.65
Lines intercon.			7-23-59	+.50	-.50
RV-15	9-10-58	-1.34 \pm .03	7-25-59	-.74 \pm .05	-.60
Lines intercon.			8-27-59	+.60	-.60
Lines intercon.*			9-12-59	+.58	-.58

These were two different samples which were interconnected and allowed to equilibrate at the end of a run by dropping an iron slug on a fragile glass septum between the two lines.

this particular zoned pegmatite can be

postulate of growth by diffusion from the rock [Ramberg 1956], when taken with known theoretical considerations [1954], seems incapable of explaining the variation in feldspar isotopic ratios at the distances observed. Also, any isotopic variation produced during diffusion of silica in a pegmatite from the outer boundaries result in enrichment of the lighter isotope at the center of the pegmatite. This trend is opposite to the observed increase in the concentration of the heavier isotope of silicon in the quartz adjacent to the central quartz core.

Another possible crude model is a closed system containing a liquid magma which solidifies from a melt, cooling and crystallizing inwardly from the outer edges, might be expected to have an outer zone displaying smaller isotopic differences between the quartz and the feldspar than the later-solidifying inner zones. If the outer zone crystallized, water would be more concentrated in the inner zones, due to the well-known depression of the solidification temperature [Truttle and Bowen, 1954], and this depression would be reflected in larger isotopic differences between the quartz

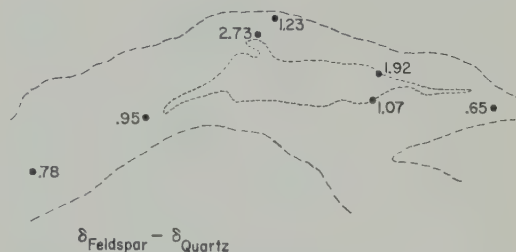


Fig. 3. Difference in silicon isotope ratios between quartz and feldspar, in per mil.

and the feldspar. Differences between the quartz and the feldspar were found to increase from the outer zones toward the core. In such a model, as the water concentration increases the temperature of solidification decreases, and isotopic exchange between the solidifying minerals therefore continues until a lower temperature is reached. It would thus be expected that in those parts last to solidify, where feldspar was found most enriched in Si^{30} , the quartz would be found correspondingly most depleted in Si^{30} . This was not observed. The quartz at points of heaviest feldspar did not differ significantly from that at points of lightest feldspar. Thus a model involving simple equilibrium solidification of a liquid melt in a closed system appears inadequate.

2. Rose Quartz Pegmatite Samples from Pala, California

δ_{feldspar}	δ_{quartz}	$\delta_{\text{feldspar}} - \delta_{\text{quartz}}$
0.85 \pm .15	0.20 \pm .15	0.65
.58 \pm .15	-.20 \pm .15	.78
.91 \pm .15	-.04 \pm .15	.95
.70 \pm .15	-.37 \pm .15	1.07
1.20 \pm .15	-.03 \pm .15	1.23
1.85 \pm .15	-.07 \pm .15	1.92
2.97 \pm .15	0.24 \pm .15	2.73

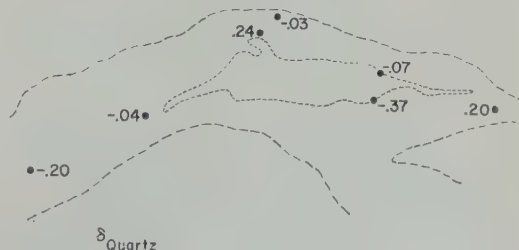


Fig. 4. Quartz δ values, in per mil.



Fig. 5. Feldspar δ values, in per mil.

to explain the observed distribution of silicon isotopes.

The postulate of an open system in which crystals grew from large amounts of percolating aqueous solution seems incapable of explaining the extreme variation of silicon ratios in the feldspar over a distance of only 20 feet, unless (1) the isotopic abundance of silicon in the percolating solutions changed very greatly or (2) the temperature at which crystallization took place from the percolating solutions changed very greatly during crystallization. Explanation 1 seems incompatible with the fact that the quartz in contact with the heaviest feldspar is approximately the same isotopically as the quartz in contact with the lightest feldspar. Explanation 2 would eliminate the above objection, but is reasonable only if the fractionation between quartz and the silica in solution was extremely small and the isotopic composition of silica in the solution remained constant. In such a case, even if the temperature of equilibrium changed enough to affect the feldspar isotope ratios markedly, the resulting shifts in the quartz ratios might be unobservable.

Considerable recent work with laboratory-grown pegmatites (R. H. Jahns, private communication) strongly suggests that parts or all of the pegmatites grow from a vapor phase or from a magma-vapor interface. If crystals grow from the vapor phase, considerable isotopic

fractionation may result. Extreme enrichment of the last crystals to solidify may be expected if the conditions necessary for Rayleigh fractionation exist. The basic condition for such fractionation is that one phase forms by transport of material from another phase.

If there is equilibrium isotopic enrichment between two phases in a closed system, the maximum differential enrichment attainable is given directly by the normal enrichment factor. If, however, one phase is continuously removed from the system or prevented from continuing reaction, as in the case of rapid crystallization from a vapor, it is possible to obtain a large isotopic enrichment of the residual material.

In idealized form the following are the conditions for observable isotopic enrichment by Rayleigh fractionation:

1. A fixed amount of vapor exists at the beginning of crystallization in the system, and no new vapor formed or added during crystallization.
2. A nonzero isotopic fractionation exists between crystals and vapor.
3. Crystals grow at a sufficient rate to prevent re-equilibration of early grown crystals with later vapor.
4. Crystallization proceeds at a sufficiently slow rate so that the vapor remains homogeneous in isotopic composition over a sizable (macroscopic) volume.
5. The vapor phase exists in an interconnected volume, in which crystal growth occurs and not merely in small regions separated by impermeable material.

Such a closed-system crystallization process may be defined as a Rayleigh crystallization process, by analogy with the more common Rayleigh distillation process.

The volume of material within which conditions 4 and 5 must be satisfied for visible isotopic enrichment to be observable will depend both on the sensitivity of measurements and on the distance between available samples. Conditions 1 and 3 may be only partially fulfilled without eliminating observable isotopic fractionation.

The basic equation describing isotopic enrichment by Rayleigh fractionation during conversion of phase 1 to phase 2 at a fixed temperature is [Rayleigh, 1896]

$$R_1(f) = R_0 f^{(1-\alpha)}$$

TABLE 3. Quartz and Feldspar Differences in Yosemite Rocks

Rock	$\delta_{\text{feldspar}} - \delta_{\text{quartz}}$
Arch Rock granite	1.5 \pm .3
Gateway granodiorite	2.75 \pm .3
El Capitan granite	1.3 \pm .3
Cathedral Peak granite	-0.2 \pm .2
Hoffman pegmatite	2.25 \pm .4

- = isotopic abundance ratio before commencement of fractionation.
- = fraction of first phase not yet removed to phase 2.
- = isotopic abundance ratio of remaining fraction f of phase 1.
- = R_1/R_2 = equilibrium fractionation factor between the two phases, a function of temperature.

define $\delta_0 = \delta$ as defined in equation 1 for fraction of phase 2 to be removed, with δ' and $\delta'(f) = \delta$ as defined in equation 1 for fraction of phase 2 formed when a fraction f of phase 1 remains, and make use of the fact $1/(1 + \epsilon) \approx 1 - \epsilon$ for $\epsilon \ll 1$, it can be shown

$$\delta'(f) - \delta_0 = 1000(1 - \alpha) \ln f \quad (3)$$

is evident from inspection that $\delta'(f)$ —the δ value of phase 2 (or crystalline material) changes markedly as f becomes sufficiently small. This is the typical behavior of a separation process.

The above equations in the form given pertain to crystallization of a single mineral from a closed vapor phase and can be quantitatively only if the listed conditions are met. It is thus difficult to apply them to a pegmatite.

With the very small number of samples (however, a region of feldspar core enriched in Si^{30} was observed adjacent to quartz core (samples RQ-3, RQ-8, and, to a lesser degree, RQ-7). A difference of 1.5 per mil was observed between feldspars from RQ-3 and RQ-7, which were only 20 feet apart. Such large shifts in isotope ratios in such small distances are difficult to explain except on the basis of vapor transport and resulting Rayleigh fractionation.

It is striking that no comparably enriched or depleted quartz samples were found. This may be the result of inadequate sampling of the pegmatite. No samples were available from the interior of the quartz core or from any other part except surface exposures. The reasons for the similarity of all the observed quartz δ values are still unclear.

A detailed discussion of the petrology, geochemistry, and crystallization of the Rose

Quartz Pegmatite will be presented separately.

Summary and conclusions. Measurements of silicon isotope ratios have been made on several samples of quartz and feldspar from a zoned pegmatite. The feldspar samples were in all cases enriched in Si^{30} with respect to quartz. The enrichment varied from about 0.6 to 2.7 per mil in the sample pairs studied. Most of the observed variation in enrichment was in the feldspar; only small variations were found in the quartz. The feldspar enrichment increased systematically from the outer edge of the pegmatite to the inner quartz core boundary.

The observations in feldspar are most readily explained by assuming crystal growth from an interconnected vapor phase, with silicon isotope fractionation between solid and vapor. The details of the formation of the quartz core and the order of crystallization cannot be ascertained from the limited observations reported. The variations in the abundances of silicon isotopes, however, are large enough and can be measured with sufficient precision to justify more complete studies of the isotopes of natural pegmatites.

Acknowledgments. I should like to thank Professor R. H. Jahns for providing the Rose Quartz Pegmatite samples and a map, and for discussing freely his unpublished recent work on pegmatites; Professor J. H. Reynolds, who directed this research; and O. Tobisch, who did all the mineral separation and sample chemistry for this study. Financial assistance was provided by the U. S. Atomic Energy Commission and the Institute of Geophysics of the University of California.

APPENDIX A

Rose Quartz Pegmatite sample locations and descriptions.

- RQ 1. Massive quartz; edge of core low on SW corner of main quartz knob. Hanging-wall edge of core.
- RQ 2. Perthite; pieces from large crystal in contact with quartz of RQ 1.
- RQ 3. Quartz and alkali feldspar; immediately east of quartz core, pit on east side of knob at top. Beneath foot-wall edge of core.
- RQ 5. Graphic granite; adjacent (on NW) to quartz core, main NW ridge. About 2 feet above hanging-wall margin of core.
- RQ 6. Graphic granite merging into small pod of quartz-perthite-muscovite pegmatite; last pit low on NW ridge. Hanging-wall

half of relatively thin part of pegmatite body that contains no quartz core.

RQ 7. Fine-grained, somewhat aplitic rock; main ridge east of quartz core. Typical of foot wall 2 to 4 feet of pegmatite body.

RQ 8. Quartz and perthite; SE edge of quartz core along top of south ridge. Foot-wall edge of core.

RQ 10. Graphic granite; pit low on south ridge. Hanging-wall half of relatively thin part of pegmatite body (some central quartz pods, but no well-defined quartz here).

REFERENCES

Grant, F. S., The geological significance of vari-

ations in the abundance of the isotopes of silicon in rocks, *Geochim. et Cosmochim. Acta*, 5, 225-231, 1954.

Ramberg, Hans, Pegmatites in West Greenland, *Bull. Geol. Soc. Am.*, 67, 185-213, 1956.

Rayleigh, J. W. S., Theoretical considerations respecting the separation of gases by diffusion and similar processes, *Phil. Mag.*, 42, 493-498, 1876.

Tilles, D., Natural variations in isotopic abundances for silicon, *J. Geophys. Research*, 66, 3033-3037, 1961.

Tuttle, O. F., and N. L. Bowen, Origin of granite in the light of experimental studies in the system $\text{NaAlSi}_3\text{O}_8\text{-KAlSi}_3\text{O}_8\text{-SiO}_2\text{-H}_2\text{O}$, *Geol. Soc. Am. Mem.* 74, 1958.

(Manuscript received November 2, 1960; revised June 7, 1961.)

Magnetic Properties of Diabase Dikes

D. W. STRANGWAY¹

*Department of Physics
University of Toronto*

Abstract. A study has been made of the magnetic properties of several diabase dikes of Cambrian age from the Canadian shield. These dikes are of widely differing ages and belong to several swarms having widely different strike directions. It was found that the direction of remanent magnetization of the various dikes was quite widely scattered before they were magnetized in a-c fields of 100 to 200 oersteds. However, after demagnetization a weak remanent magnetization was found whose direction was stable in a-c fields up to 500 oe. In most all the cases studied the direction of this stable magnetization was found to lie close to the plane of the dike, although the various dike swarms had different inclinations within the plane. It was shown experimentally that typical samples acquired a large, soft TRM at 585°C when cooled in the earth's field. A second weak but very stable component was also found which had a Curie temperature considerably less than 585°C. It is postulated that the component with the high Curie temperature produced internal demagnetization which tended to cancel out the component of the earth's magnetic field normal to the dike. The component with the lower Curie temperature then acquired a magnetic moment parallel to the internal field of the dike, which would be nearly in the plane of the dike. This stable component may have been preserved since the time of formation of the dike.

Introduction. In the Canadian shield there are several swarms of diabase dikes which have been known to geologists. However, wide-aeromagnetic coverage is revealing the existence of a far greater number of these than has been suspected previously. Moreover, the spinner magnetometer is proving that some of these dikes are as much as 300 miles long, although only a few hundred feet wide. The dikes are all believed to be of Precambrian or very early Cambrian age, but there are known to be wide differences in age. The individual swarms have different strike directions and have been classified separately as, for example, the Matachewan, Grenville, and Grenville types. (For a more detailed discussion of the dike classification, see Strangway [1960].)

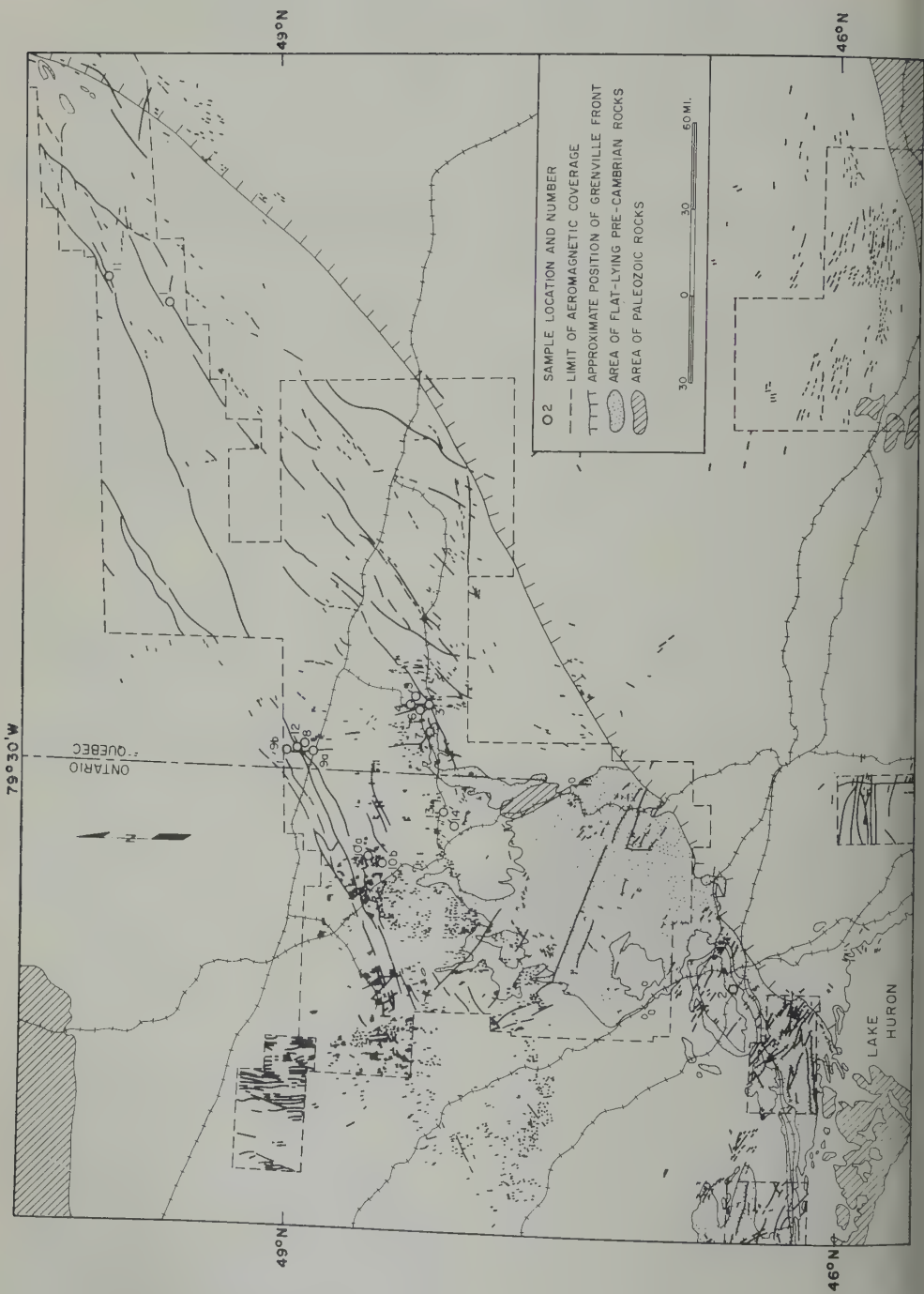
Because of the geological importance of these dikes it was decided that a study of the magnetic properties might be useful in throwing light on their relative ages as well as on the general properties of dikes in general.

Sample collection and measurement. One hundred and eighty-six samples were collected from the dikes shown in the accompanying general-

ized map of portions of Ontario and Quebec (Fig. 1). Samples were collected from 19 dikes at 14 different sites. Samples from three of the dikes had remanent magnetizations below the limit of measurement (5×10^{-5} emu/cm³). At each dike approximately ten samples were taken to test for variation in properties across the width of the dike. Each sample was cut into several 1-inch cubes and the direction and intensity of the remanent magnetization were measured. This was done with a spinner magnetometer [Hood, 1956] in which the cubes were spun at 40 rev/sec adjacent to a pickup coil. The reversible susceptibility of each sample was also measured in a 40-cps a-c field with a peak intensity of about 1.0 oe.

Fourteen of the 19 dikes were examined in detail. None showed any regular variation in the value of susceptibility (k), in the value of remanent magnetization (J_r), or in the ratio of the remanent to the induced magnetization ($Q = J_r/kH$) when the values across the width of the dike were compared. This was somewhat surprising in view of the observed large gradation of grain size from the edges to the middle of any one dike. This may be due to the fact (as noted in thin sections and polished sections) that the coarse-grained portions of the dikes

¹Present address: Department of Geology, University of Colorado, Boulder.



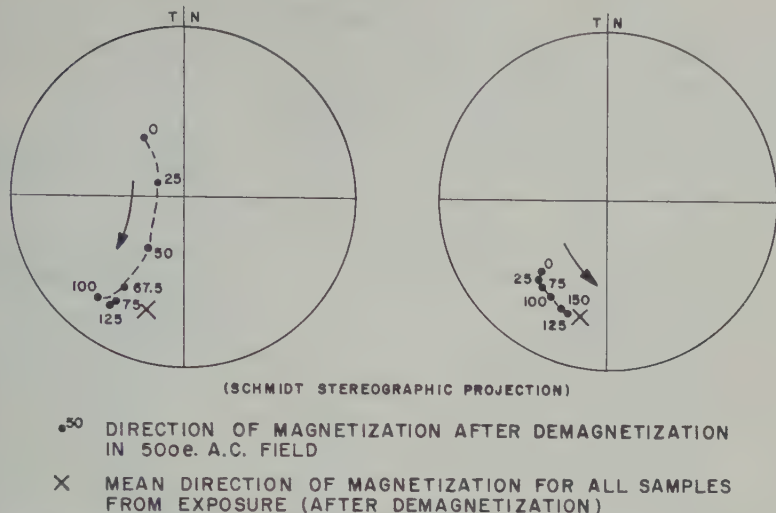


Fig. 2. Change in direction of remanent magnetization with increasing a-c demagnetization field.

ned intimate intergrowths of magnetite and ilmenite while the fine-grained portions had discrete particles. Consequently, the magnetite and ilmenite particles in the center of the dike had about the same dimensions as individual grains at the edge of the dike.

Because the initial directions of magnetization were usually found to be quite widely scattered, at least one cube from every sample was subjected to a-c demagnetization. The a-c (60 cps) demagnetization was carried out at various field strengths in a space that was free of d-c field (within about 5 per cent). The purpose of demagnetizing most of the earth's field was to avoid spurious 'anhysteretic' magnetization described by Rimbert [1957] and others. The problem of 'anhysteretic' magnetization proved very serious because of the large soft components of magnetization that were encountered. It was found that by demagnetizing a second time after rotating the sample through 180°, that any residual d-c field acted in opposite directions, it was possible to detect the presence of a spurious magnetization. If after demagnetization the two magnetization vectors were found to differ by less than 60°, a simple averaging process reproduced the true direction of stable magnetization with considerable accuracy. This double demagnetization was repeated along each of the three cube axes and the direction of stable magnetization determined.

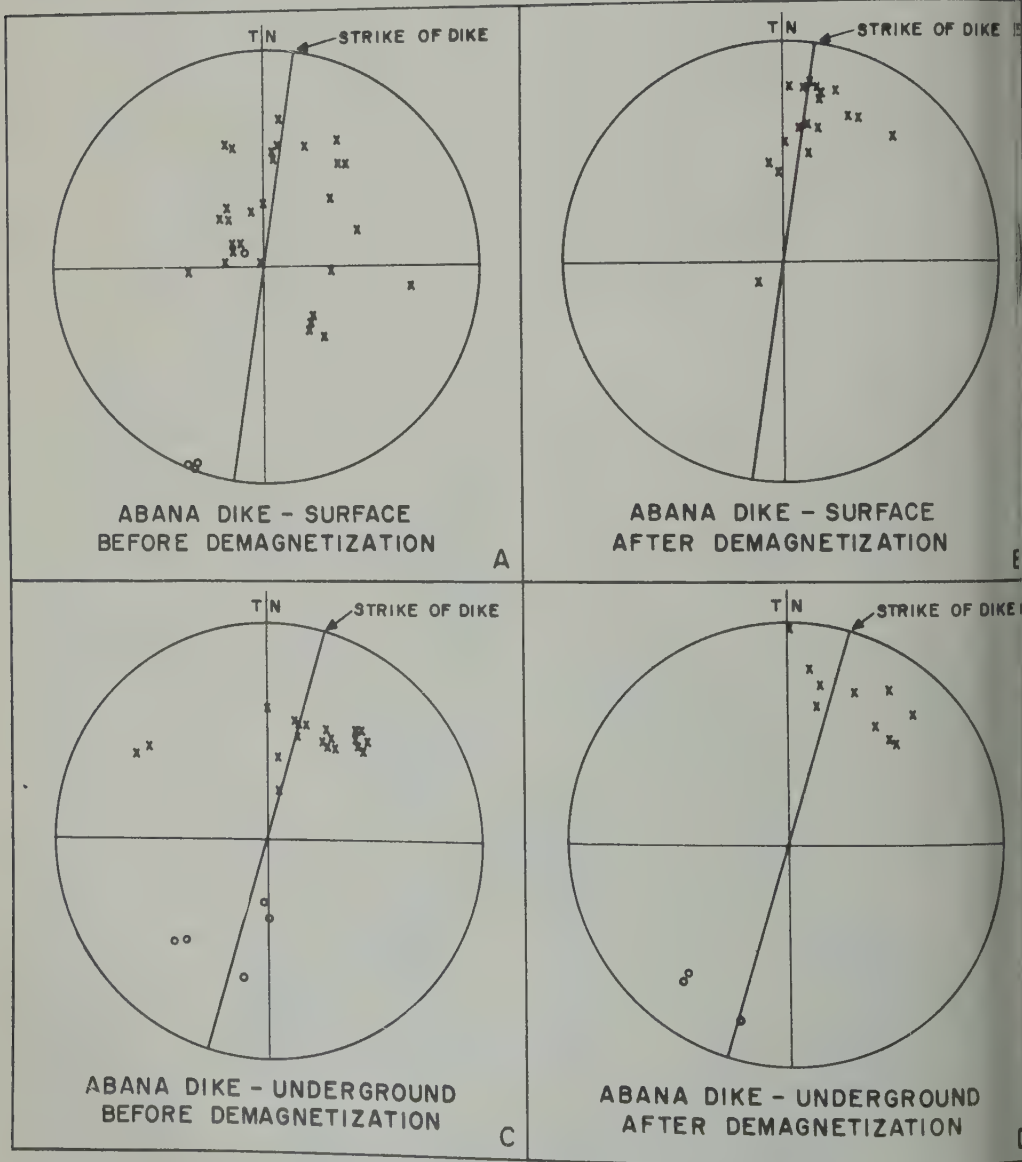
For a few samples from each dike this process was carried out using a-c fields of increasing strength. It was found generally that at a-c fields of 100 to 200 oe, the direction of magnetization had become stable to increases in field strength. In two of the dikes no magnetization was detectable after demagnetization in fields over 50 oe, probably indicating that the stable magnetization found in most of the dikes was not present in these two. Figure 2 shows two examples of how the direction of magnetization approached a final value. After the minimum field necessary to reach stable magnetization was determined for any given dike, at least one, and usually two, cubes from every sample was demagnetized using this minimum field strength. A second measurement using a higher field strength was usually carried out to confirm the direction of stable magnetization.

Figure 3 shows data for two exposures from one dike. It can be seen that the grouping of the results is much better after than before demagnetization, which lends support to the idea that a soft component of magnetization having a random direction has been removed by the demagnetization process. This same improvement in grouping after demagnetization was observed in most of the exposures studied.

One of the major unexplained problems in rock magnetism is that of magnetic reversals. In Figure 3D three cubes show a stable magnet-

ization exactly reversed to that of the bulk of the samples. It is difficult to explain these results except on the basis of some still unknown physico-chemical process. If the explanation were based on a reversal of the magnetic field during cooling, one would expect the reversed

samples to be symmetrically placed on the two sides of the dike. However, these samples came from one edge of the dike while the other edge showed only a normal polarization. It is suggested that some unknown physico-chemical process has caused this reversal (and several



N. SEEKING POLE IN UPPER HEMISPHERE
N. SEEKING POLE IN LOWER HEMISPHERE

Fig. 3. Stereographic plots of directions of remanent magnetization showing effect of demagnetizing.

TABLE 1. Summary of Paleomagnetic Results after Demagnetization

No.	Type of Dike	Strike of Dike, deg	Avg.		No. of Samples	α , deg	K
			Declination of J_r , deg	Inclination of J_r , deg			
	ENE dike parallel to Grenville Front	258	263	-56	9	10.8	23.5
	NW Sudbury olivine dike	297	266	-8	8	25.0	5.1
	N-S dike (Noranda area)	256	255	-61	7	15.9	12.8
	N-S dike (Noranda area) (Matachewan type)	191	215	+8	8	25.9	4.9
	E-W dike (Noranda area)	226	246	-47	5	18.0	12.9
	N-S dike (Noranda area)	164	161	+19	8	16.1	12.5
	N-W dike (Noranda area)	154	190	+10	5	40.5	1.2
	ENE dike parallel to Grenville Front	251	266	-42	13	8.8	24.5
	N-S dike (Normetal area) (Abana dike)	16	18	-30	31	5.3	24.2
	N-S Matachewan dikes	176 & 197	186	+3	21	19.2	6.5
	ENE dike parallel to Grenville Front	243	233	-69	1		
14	(Remanent magnetic intensities less than 5×10^{-5} emu/cm ³)						

ones found in this study), and the pre-succussions have been made assuming all to be normally magnetized.

When the direction of stable remagnetization of each sample was plotted, it was observed that the mean direction of magnetization was always close to the plane of the dike involved. The inclination within the plane varied from one dike swarm to another, but seemed to be a strong correlation with the strike in each case. This is shown in Figures 3 and Table 1. Those specimens which had very high Q values were rejected, as they undoubtedly remagnetized by lightning. However, the rest of the samples which had been demagnetized were analyzed statistically using the Fisher [1953] analysis for directions on a sphere. As is commonly done in paleomagnetism, a 5 per cent limit of confidence was placed on the results so that there is a 95 per cent chance that the true mean lies within α of the calculated mean. Table 1 includes the results of these computations for the dikes studied and also gives the strike direction and inclination from which the samples were obtained. The value K is a precision parameter which gives an indication of the dispersion of directions of magnetization. When $K = 1$, the directions are completely randomly distributed, and as K becomes larger the grouping becomes more pronounced.

With the exception of dikes 2 and 7, the directions of the average declinations lie within 24° , and in most cases within 15° , of the strike directions of the respective dikes. In the case of dikes 2 and 7, no stable component of magnetization was found, since the intensity of magnetization became extremely small and unmeasurable before a stable direction could be reached. As will be pointed out later, it is possible that the applied field may have been nearly normal to the plane of these dikes at the time they were last cooled through their Curie temperatures. It is interesting to note that dikes 2 and 7 both trend toward the northwest, and they apparently bear no relation to any of the other dike swarms studied.

The relation between strike direction and direction of remanent magnetization has been summarized on an equal-area stereographic contour plot. In Figure 4 the results of 132 demagnetization measurements have been compiled with the strike of the respective dikes as the reference plane. Contouring is carried out in the usual fashion (1 per cent of the total area of the projection contains a given percentage of the total number of points present). In this plot it is evident that up to 12 per cent of the total number of points are contained within 1 per cent of the area. The correlation of the remanent magnetization vector with the strike of the dike is very obvious. No contour

intervals are located more than about 20° from the plane defined by the dikes.

There have been a few other studies of similar dikes with a range of ages and in various parts of the world. Among these studies are those by *Graham* [1953] on dikes in northern Michigan, by *Bruckshaw and Robertson* [1948] on dikes in England and Scotland, and by *Gough* [1956] on dikes in South Africa. The samples in most of these studies have not been subjected to routine demagnetization, but fairly consistent results have nevertheless been found. In the present study it was found that when the initial results were well grouped, the demagnetization of the samples did not change the mean direction a great deal. All the results obtained in the present study and all those by the above authors have been compiled in Figure 5 on an equal-area contour map. Once again, there is a very definite correlation between the direction of stable magnetization and the strike of the respective dike. The concentration of points at the center of the projection is the result of the large number of samples collected by *Gough* of a few dikes of the same age.

Probable physical explanation. The evidence presented here suggests that the most stable component of magnetization in diabase dikes lies close to the plane of the dike. Much work in the field of paleomagnetism assumes *a priori*

that the direction of magnetization acquired and retained throughout the lifetime of a sample is parallel to the earth's field at the time of magnetization. However, it seems most unlikely that the direction of the external applied field would always happen to lie parallel to the strike of dikes of several different ages and with different geographic distribution.

Several possible explanations of this anisotropy have been considered. It seems unlikely that there is any preferred grain orientation in the dikes, hence any possible magnetic anisotropy in the dikes. The dikes are all gradational from edge to middle with fine-grained material at the edge of the dike becoming coarse-grained in the middle. This is consistent with a cooling process during which crystallization took place after emplacement. No flow structures have been observed. Thin sections indicate that a random directional distribution of grains exists both at the edge and in the center of the dikes.

Pressure effects cannot be eliminated as a possible explanation of the results, but pressure causes seem unlikely. If pressure effects were responsible for the magnetic orientation, it is probable that all samples would be magnetized in the same direction (either horizontally or vertically) which is contrary to the large range of inclinations found in this study.

The most plausible explanation appears

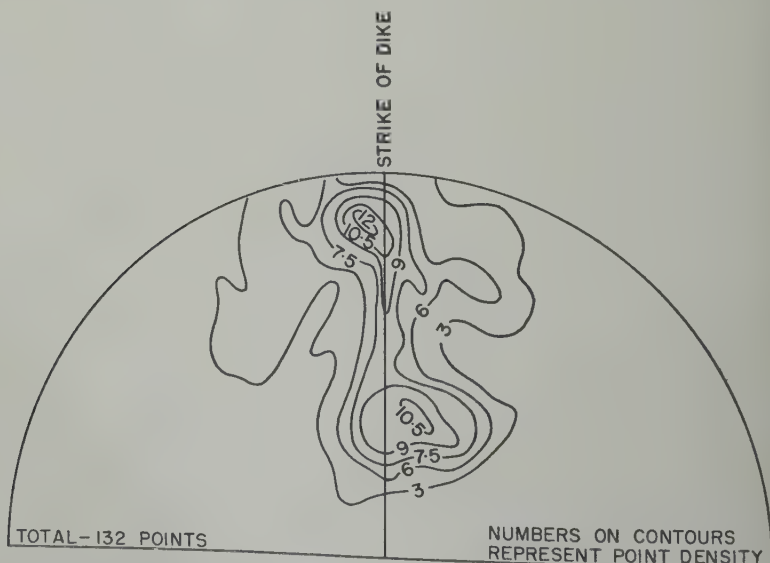
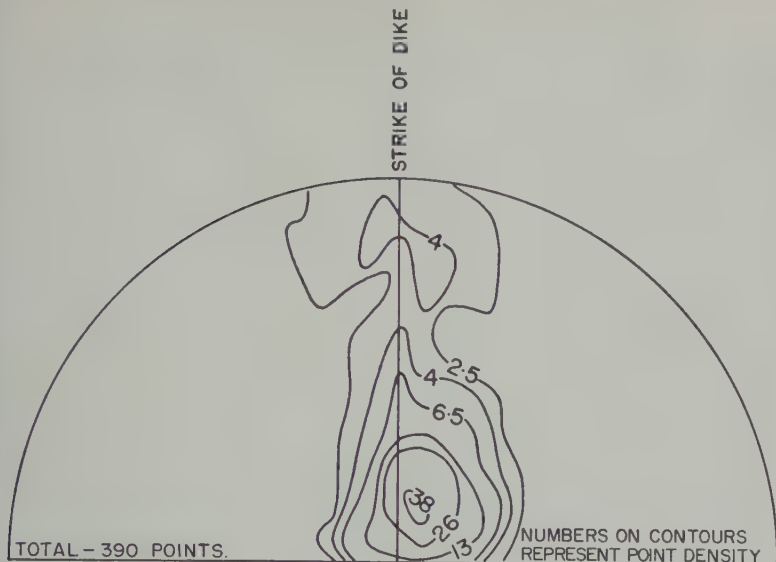


Fig. 4. Stereographic plot of directions of remanent magnetism in dikes from N.E. Ontario and N.W. Quebec showing relationship to strike of the dike.



5. Stereographic plot of directions of remanent magnetism in dikes for which data have been published showing relationship to strike of the dike.

based on demagnetization within an infinite sheet. In Figure 6 we see the various fields acting on any particular point within a dike. There is the external field of the earth, within a magnetic sheet there is in addition a demagnetizing field H_{dem} , a Lorentz field and a field due to the adjacent particles. This magnetizing field is due to the shape of the dike and it is normal to the infinite sheet. It opposes the external field and so tends to cancel the normal component of the external field. The Lorentz field is a magnetizing field due to the magnetization. This field is taken into account by setting up within a spherical cavity with radius r in consideration at the center of the dike. The size of this cavity is taken to be large enough that individual magnetic grains within the cavity exert a field equivalent to a uniform magnetization. Finally, it is necessary to add the effect of the individual grains within the Lorentz sphere. If these are randomly distributed, the most probable field within the sphere due to these grains is zero. The internal field H_{int} , then, is

$$H_{int} = H_{ext} - H_{dem} + H_{Lor}$$

Components H_{par} parallel and H_{perp} perpendicular to the dike, we have

$$H_{par} = H_{ext} \cos \alpha + \frac{4}{3}\pi M \cos \alpha$$

$$H_{perp} = H_{ext} \sin \alpha + \frac{4}{3}\pi M \sin \alpha - 4\pi M \sin \alpha$$

where M is the total magnetization $= (kH_{int} + J_r)$ and $k =$ vol. susceptibility.

Since the susceptibility magnetization is dependent on the internal field, this latter magnetization is analogous with a refraction process. From Figure 7 it is apparent that the susceptibility of the diabase dikes does not change the direction of the internal field by more than a few degrees. However, in the same figure a cal-

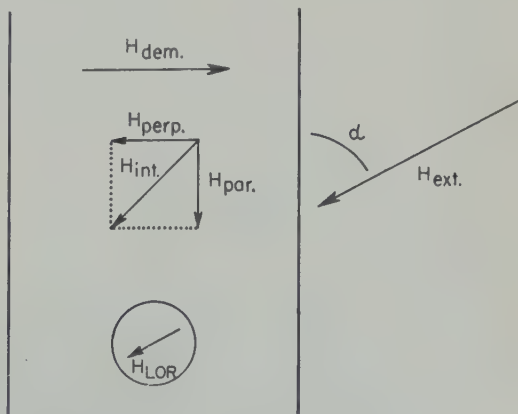


Fig. 6. Fields acting within a magnetic sheet.

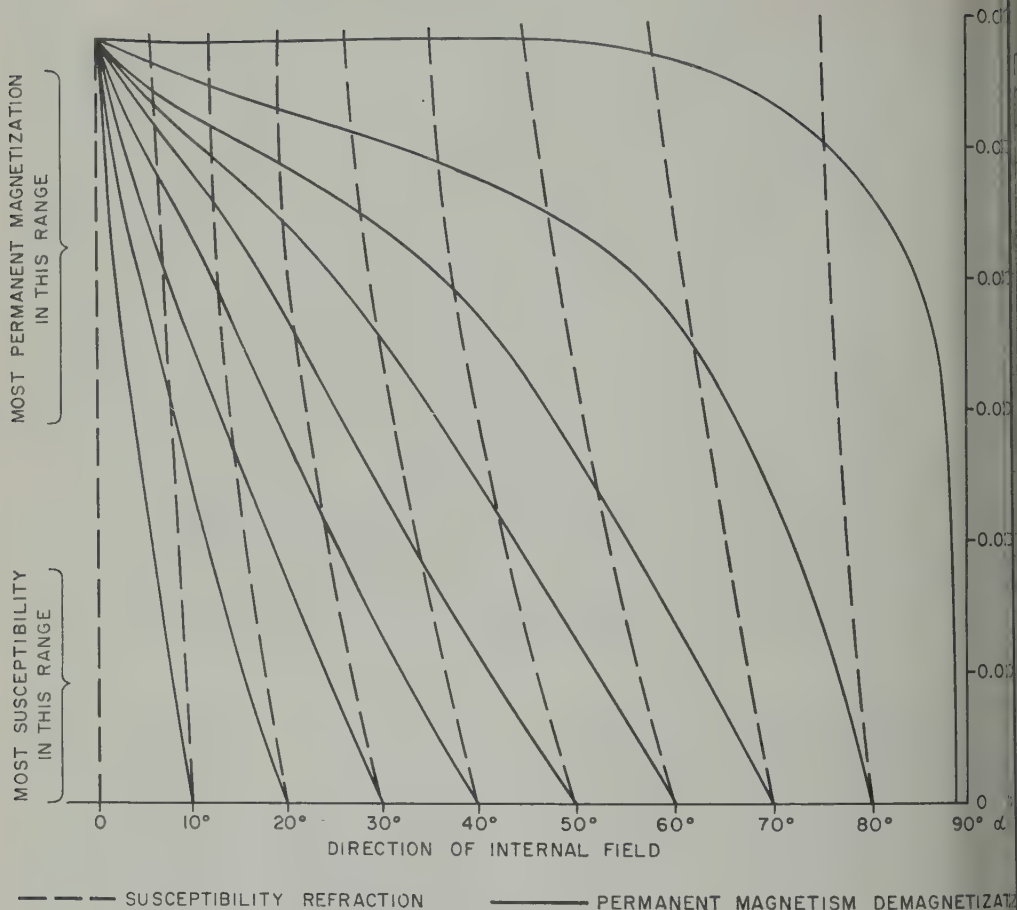


Fig. 7. Variation of direction of internal field with various directions of applied field (α) and various intensities of susceptibility and permanent magnetization.

culuation of the direction of the internal field, account having been taken of the demagnetizing and Lorentz fields, shows that the direction of the internal field can be markedly different from that of the external field if a sufficiently large permanent magnetization is present.

If this internal demagnetization is to explain the observed experimental results, a number of conditions must be satisfied:

1. The intensity of freshly acquired remanent magnetism must be relatively large (0.02 to 0.05 emu/cm³).

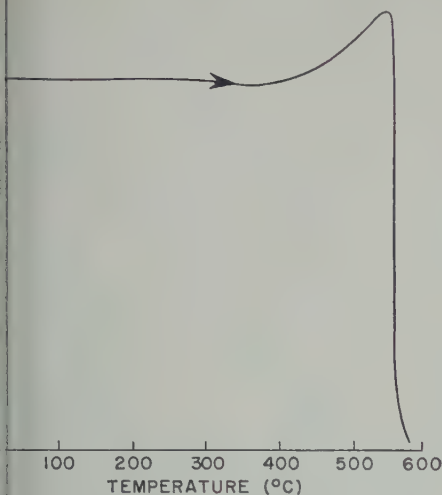
2. The main body of magnetization must be acquired at a single Curie temperature and be parallel to the applied field, H_{ext} .

3. This magnetization must be sufficiently stable to remain in a fixed direction as the body cools; otherwise the demagnetization effect

would be slight. However, this magnetization must be unstable over time intervals of the order of the age of the rock to explain the scattered results found today.

4. Finally, a second more stable component with a lower Curie point must also be present.

As a check on whether these conditions are satisfied, laboratory experiments were carried out on 20 representative samples from the edges and centers of the dikes. The samples were heated above the highest Curie temperature and allowed to cool in the present earth's field. The intensities of remanent magnetization in the cooled samples varied from about 0.015 to 0.04 emu/cm³, with most of the values falling about 0.03 to 0.04 emu/cm³. The intensity of magnetization after heating was generally increased by a factor of 7 or more, indicating



Typical plot of susceptibility versus temperature (readings taken at approximately 10°C intervals).

great deal of the original TRM has decayed in geologic time (or that the earth's field is now 10 times less than it is at present). The intensity of magnetization would certainly be sufficient to reduce greatly the normal component of the internal field of the dike.

On heating these samples, measurements of susceptibility were made at intervals of about

7°C, and almost without exception the curves were typical of a single magnetic component with a Curie temperature of about 580°C (see example in Fig. 8). This suggested that the main component of magnetization was due to magnetite. While most of these samples were being cooled in the earth's field they were rotated at about 5°C below this Curie temperature; therefore their orientation with respect to the applied earth's field was different for the remainder of the cooling process. The direction of magnetization of the samples immediately after cooling was invariably found to be parallel to the original field direction, indicating that the major component of magnetization was determined by the direction of the applied field at the Curie temperature. In other words, the net magnetization of dikes just after cooling is parallel to the external field and is sufficiently stable not to change direction during the cooling process. The main portion of the magnetization is therefore available to set up internal demagnetization.

Following the measurement of the magnitude and direction of the remanent magnetization produced by the cooling process, the samples were subjected to successive a-c demagnetization. It was found that the intensity of magnetization dropped off rapidly, and the direc-

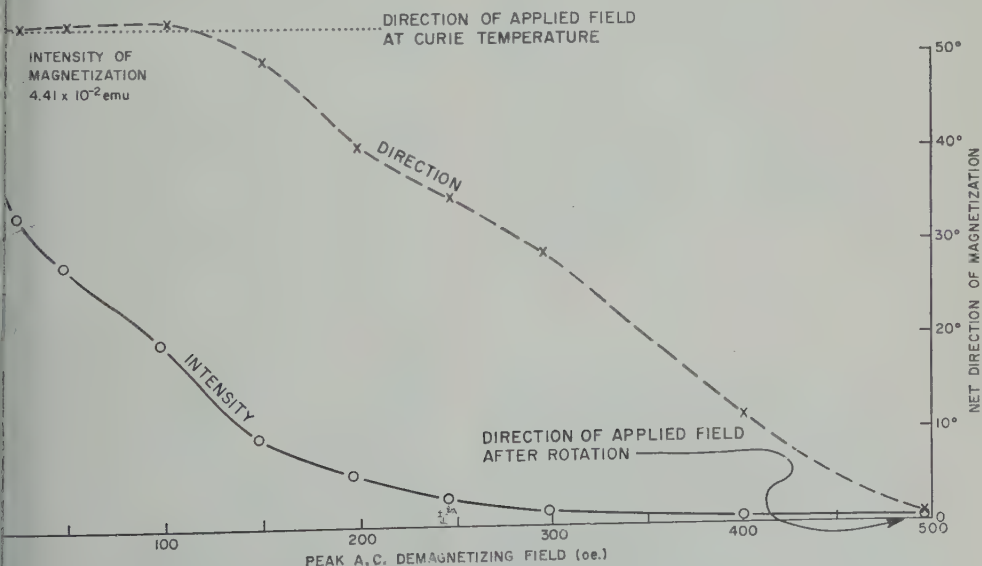


Fig. 9. Results of successive a-c demagnetization of a sample which has been cooled through the Curie temperature in a 0.5-oe field and rotated through a 50° angle at 5°C below Curie temperature.

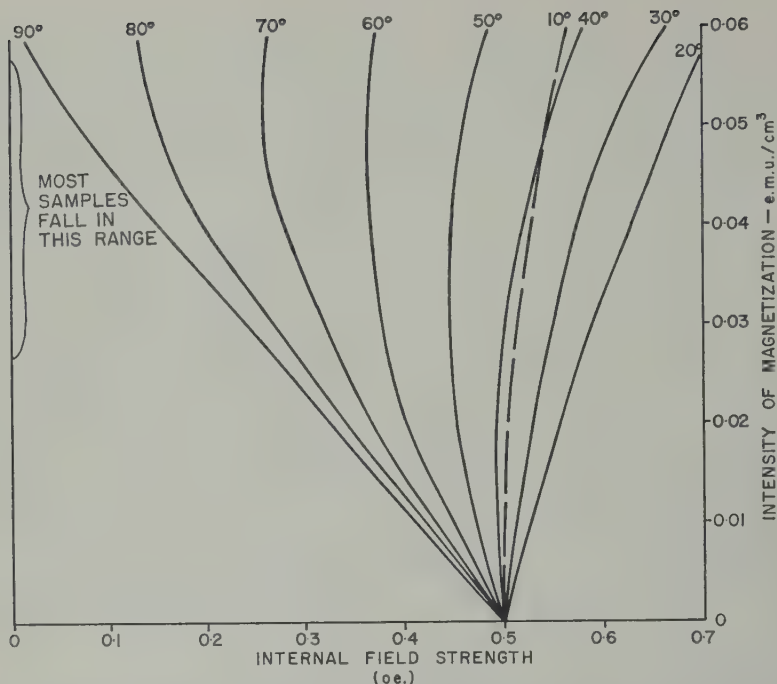


Fig. 10. Variation of internal field strength for various intensities of magnetization and various directions of permanent magnetization (including Lorentz field).

tion of magnetization changed gradually until it approached that of the second applied field. This is illustrated for a typical sample in Figure 9. This experiment established the presence of a stable, low-intensity component of magnetization, having a lower Curie temperature and much greater stability than the main component.

It is postulated that this is the magnetization whose direction is preserved throughout the lifetime of the diabase. Some of the other components may also be preserved, but it seems likely that the softer component will decay more rapidly with time and be more subject to alteration.

The model may also provide an explanation of certain other aspects of the natural magnetization process. If the applied external field is originally normal to the dike, the intensity of the internal field becomes very small (see Fig. 10). Nagata [1953] has shown that the intensity of magnetization is proportional to the applied field. As a result, in this case the stable component of magnetization will be very small. This possibility may explain the results obtained for dikes 2 and 7, where no measurable stable

component was found. When samples from these dikes were allowed to cool through the Curie temperatures, stable components of magnetization were produced indicating that the magnetic properties are similar to those of other dikes.

Source of stable component. The rocks were examined for possible sources of the hard component of magnetization. About 30 thin sections and 30 polished sections were studied to obtain some idea of the distribution of the iron oxides. In general, the samples contained 5 to 10 percent of these minerals. In coarse-grained samples, large grains (about 0.5 mm) of the minerals were found as intimate sheetlike intergrowths of magnetite and ilmenite (see Fig. 1). The Curie point measurements suggest that magnetite is almost always present as a discrete phase. If it were in solid solution with ilmenite the Curie point would be lowered. The laths of magnetite and ilmenite are probably only a few microns thick, although this dimension is difficult to estimate since most of the polished sections are not cut normal to the intergrowth plane. This conclusion is supported by the fact that the fine-grained edge samples with no

growths have almost identical magnetic properties to the coarse-grained center samples. Distributed in the coarse-grained portions are finer-grained iron oxides having dimensions of the order of a few microns were also found. These grains may be the low Curie point component. However, it seems unlikely that this is the source, as all the grains in the edge samples have approximately these same dimensions and still have the same magnetic properties. In study of the thin sections, however, another distribution of opaque mineral believed to be magnetite was found, occurring in the olivine and pyroxene grains. This is a powder distribution of magnetite, apparently formed by reaction of the two host minerals with residual iron-rich solutions in the hydrothermal stage of rock formation. The powder magnetite grains are less than a micron and hence close to theoretical single domains in size. Néel's [1955] study of single-domain grains suggests that grains having dimensions of about 0.05μ will have remanent magnetizations of great stability. It is suggested that this powder magnetite in olivine and pyroxene (which forms 50 per cent of the rock) may be the site of the stable component of magnetization. In this same study Néel describes the decrease of Curie temperature with decreasing grain size. The relation he describes is based on relaxation theory in which thermal fluctuations cause destruction of the individual domains. It can be shown [Strangway, 1960], using typical values for magnetite, that for a single domain particle to have a Curie temperature of 500°C and for its magnetic moment to be stable for 10^9 years, the grain must be of the order of 0.05μ . For these reasons, then, it seems probable that the stable component of magnetization in this study is due to the powder magnetite formed by reaction within pyroxene and olivine crystals.

Conclusions. It can be stated that a correlation between stable remanent magnetic direction and the strike of diabase dikes has been established. This correlation seems well defined, and a model to explain this anisotropy has been proposed. The model depends on the demagnetization due to the shape of a dike and requires that a small portion of the magnetic mineral have a lower Curie temperature and be less stable than the bulk of the magnetic



Fig. 11. Photograph in reflected light showing lath structure of iron oxides (width of figure corresponds to 900 microns of specimen).

mineral. The presence of such a component has been proved, but it has not been possible to determine which fraction of the iron oxides is the source of this stable magnetization. It is suggested that magnetite formed by the reaction of the pyroxenes and olivines with late-stage hydrothermal iron-rich solutions is the seat of this stable magnetization.

Acknowledgments. I am grateful to R. M. Farquhar, G. F. West, and M. G. Rochester for helpful discussion in connection with this work. I also acknowledge with thanks the receipt of a Garnet W. McKee-Lachlan Gilchrist scholarship for two academic years. The project was supported by a grant from the National Research Council of Canada. I am grateful for permission to collect samples at Normetal Mining Corp. Ltd. and Teck Hughes Gold Mines Ltd. I wish to acknowledge assistance from Bear Creek Mining Company in the preparation of the manuscript.

REFERENCES

- Bruckshaw, J. M., and E. I. Robertson, The magnetic properties of the tholeiite dykes of north England, *Monthly Notices Roy. Astron. Soc. Geophys. Suppl.*, 5, 308-320, 1949.
- Fisher, R. A., Dispersion on a sphere, *Proc. Roy. Soc. London A*, 217, 295-305, 1953.
- Gough, D. I., A study of the paleomagnetism of the Pilansberg dikes, *Monthly Notices Roy. Astron. Soc. Geophys. Suppl.*, 7, 196-213, 1956.
- Graham, J. W., Changes of ferromagnetic minerals and their bearing on the magnetic properties of rocks, *J. Geophys. Research*, 58, 243-260, 1953.
- Hood, P. J., The design, construction and calibration

- tion of a remanent magnetometer, M. A. thesis, University of Toronto, 1956.
- Nagata, T., *Rock Magnetism*, Maruzen and Co., Tokyo, 255 pp., 1953.
- Néel, Louis, Some theoretical aspects of rock magnetism, *Advances in Phys. (Suppl., Phil. Mag.)*, 4(14), 191-243, 1955.
- Rimbert, F., Sur l'aimantation rémanente anhystérétique des ferrimagnétiques, *Compt.* 245, 406, 1957.
- Strangway, D. W., Magnetic properties of Canadian diabase dike swarms, Ph.D. University of Toronto, 1960.
- (Manuscript received May 1, 1961; rev June 26, 1961.)

The Behavior of Volatiles on the Lunar Surface¹

KENNETH WATSON, BRUCE C. MURRAY, AND HARRISON BROWN

*Division of Geological Sciences, California Institute of Technology
Pasadena, California*

Abstract. Volatiles, and water in particular, have been thought to be unstable on the lunar surface because of the rapid removal of constituents of the lunar atmosphere by solar radiation, solar wind, and gravitational escape. The limiting factor in removal of a volatile from the moon, however, is actually the evaporation rate of the solid phase, which will be collected at the coldest points on the lunar surface. We present a detailed theory of the behavior of volatiles on the lunar surface based on solid-vapor kinetic relationships, and show that water is far more stable there than the noble gases or other possible constituents of the lunar atmosphere. Numerical calculations indicate the amount of water lost from the moon since the present surface conditions were initiated is only a few grams per square centimeter of the lunar surface. The amount of ice eventually detected in lunar 'cold traps' thus will provide a sensitive indication of the degree of chemical differentiation of the moon.

1. INTRODUCTION

Previous analyses of the behavior of volatiles on the lunar surface [Spitzer, 1952; Kuiper, 1952; Miller, 1952; Öpik and Singer, 1960; Vestine, 1960] have all indicated that volatiles could not survive there for extended periods of time, and that water is particularly unstable because of its low molecular weight and ease of ionization. These investigations did not, however, recognize the amount of any volatile in the vapor phase on the moon—and hence its mass removal from the moon—is determined by the temperature of the solid phase at the coldest place on the lunar surface. This was recently pointed out by the authors [Watson, Murray, and Brown, 1961]. We have now developed a detailed theory of the behavior of volatiles on the lunar surface which takes vapor pressure equilibrium into account, and from it we show that water is actually by far the most stable of the naturally occurring volatile substances that might conceivably have been released at some time on the lunar surface. We show further that the total amount of water that could have been removed since the time the moon's surface conditions first evolved to essentially their present characteristics is quite small. Thus the present amount of lunar ice must bear a close relationship to the total amount of water ever present

on the lunar surface over the above time period. Conversely, the absence of lunar ice would indicate an extremely small amount of chemical differentiation of the lunar mass.

We will first develop the theory of the behavior of volatiles on the lunar surface, and then investigate the best numerical values of the parameters of our model to use in numerical calculations.

2. THEORY OF MIGRATION AND TRAPPING OF VOLATILES AND THEIR ESCAPE FROM THE LUNAR SURFACE

In this section we shall develop a theory of the behavior of volatile substances on the lunar surface which is based on two premises: (1) that the lunar atmosphere is so rarefied that molecular transport in the vapor phase can be described purely in terms of dynamical trajectories; and (2) that there are permanently shaded areas (cold traps) with temperatures at least as low as 120°K. It will be convenient to discuss first the steady loss from the cold traps, and, second, losses of any newly liberated volatiles during migration to the cold traps. By combining these results we will derive the equation governing the total mass removal rates of volatiles from the lunar surface.

2.1. Steady-state loss of volatiles from the cold trap. Once any substance condenses in a permanently shaded area, it will be subject to a continual evaporation loss, which may be

¹Contribution 1048, Division of Geological Sciences, California Institute of Technology.

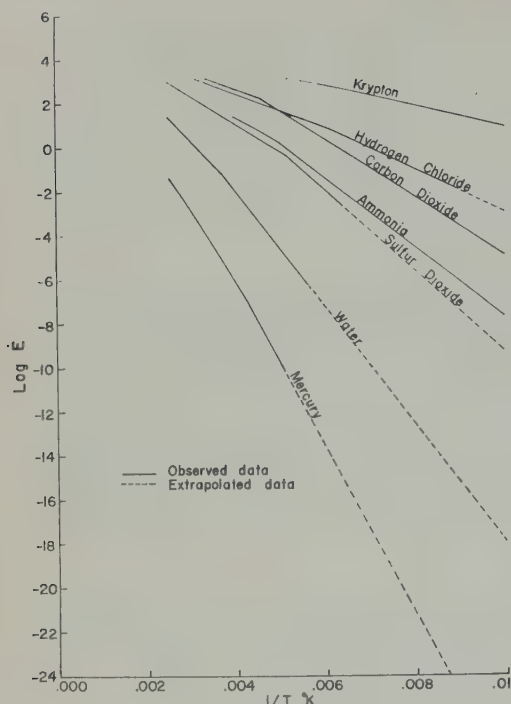


Fig. 1. Evaporation rates.

balanced in part by condensation back into the cold trap from the atmosphere. The mass loss rate from the cold traps, \dot{m}_s , is simply the difference between evaporation and condensation rates:

$$\dot{m}_s = KA(\dot{E} - \dot{C}) \quad (1)$$

where

K = fraction of lunar surface that is permanently shaded.

A = total lunar surface area in square centimeters.

\dot{E} = evaporation rate of the substance at the surface temperature of the cold trap, taken to be 120°K as an upper limit for the moon in grams per square centimeter per second.

\dot{C} = condensation rate of the substance in question back into cold trap areas in grams per square centimeter per second.

The fraction of the lunar surface in permanent shadow, K , will be estimated later in this paper to be about 5×10^{-3} . The evaporation rate is a characteristic physical property of any solid

phase that depends only on the surface temperature of that solid. The evaporation could be observed directly if the vapor removed as fast as it formed, i.e., by eliminating condensation. On the other hand, evaporation and condensation rates are equal when the vapor is saturated, if the accommodation coefficient is unity. Accordingly, the equilibrium vapor pressure present over a solid surface can be used to estimate the characteristic evaporation rate at that temperature. The maximum evaporation rate can therefore be expressed as follows [Estermann, 1955].

$$\dot{E} = p \sqrt{\frac{\mu}{2\pi RT}} = 4.374 \times 10^{-5} \times p \sqrt{\frac{\mu}{T}}$$

where

\dot{E} = mass loss in grams per square centimeter per second.

p = vapor pressure in dynes per square centimeter.

μ = molecular weight.

T = absolute temperature.

R = gas constant.

In order to obtain values of \dot{E} in the vicinity of 120°K for a number of naturally occurring volatile substances, the appropriate equilibrium vapor pressures have been extrapolated to these temperatures for the present analysis [International Critical Tables, 1928]. The experimental data are of the form $\log p = -A/T + B$, and A and B are constants characteristic of the substance, and linear extrapolation is reasonable for our purposes, assuming that no solid phase changes occur at these pressures. For ice, this assumption is experimentally validated by the work of Bridgman [1957].

The resulting maximum evaporation rates for various volatiles as a function of temperature are plotted in Figure 1.

The condensation rate \dot{C} , unlike K , A , and \dot{E} , is not a simple characteristic property of either the moon or of a particular substance because it also depends on the escape rate from the lunar atmosphere and the accommodation coefficient in the cold traps. However, it is possible to treat the migration of molecules on the lunar surface in terms of a simple probability model, provided that the average trajectory jump length of the molecules is comparable to or larger than the average size and separation of the cold traps.

derivation of \dot{C} in this way is presented now. We will demonstrate that the moon does not satisfy the restriction regarding average jump length vs. size and separation of cold traps. The following parameters will be used:

- = jump time, which is a function of the velocity distribution at the surface as well as of the molecular weight and the lunar surface gravity. \bar{t} and \bar{D} are mean jump time and mean jump length, respectively.
- = probability for a molecule to escape from the lunar atmosphere on a single jump.
- = fraction of the lunar surface occupied by permanently shaded areas.
- = accommodation coefficient, the probability that an incident molecule will stick to the surface.

The processes by which the molecules transport themselves through the rarefied atmosphere and either escape or become trapped in the permanently shaded areas are complicated in detail. For our purposes it is sufficient to realize that the fraction of the molecules that condense on a surface, as opposed to the fraction that are reflected, is controlled by the microscopic roughness and chemical nature of the surface, by the surface temperature, and by the molecular weight of the impacting molecules. A discussion of this can be found in *Loeb* [1927], *Langmuir* [1937], and *Knudsen* [1910]. We shall assume that the temperature and microscopic roughness of the cold-trap surfaces are sufficient to ensure that the accommodation coefficient is close to unity. We will, however, consider the effect of changing α when numerical calculations for the trapped fraction are performed. It is also reasonable from the above to assume that molecules landing from the warm surfaces will approach surface temperature after a few jumps, regardless of the velocity distribution at the surface.

It is apparent that those molecules captured in temporarily shaded areas, such as the dark craters, cannot escape until the host area heats up. Therefore, we can neglect the effect of these temporarily shaded areas in the solution of the problem, and consider only those molecules continually moving at any instant of time.

We shall further assume, to avoid unnecessary complexity, that the reflection and emission of

molecules from the lunar surface can be adequately described by a cosine distribution. The average jump time \bar{t} , and the average jump distance \bar{D} , are, therefore:

$$\bar{t} = 2 \times \frac{.866 \times V}{g} \quad (3)$$

$$\bar{D} = \frac{1}{2} V \bar{t} \quad (4)$$

where

V = rms velocity = $\sqrt{3RT/\mu}$.

g = surface gravitational acceleration.

R = gas constant.

T = effective surface temperature ($^{\circ}\text{K}$).

μ = molecular weight of volatile.

We now examine the implications of our assumptions. If \bar{D} is at least comparable to the size and separation of the permanently shaded areas, we can presume that for any individual jump the probability of landing on a permanently shaded area is proportional only to the total fractional area of permanent shade K , and does not depend on the location of the beginning point of the jump. Since α represents the escape probability during an individual jump, the probability of a particle landing on a cold-trap surface and not escaping is $K(1 - \alpha)$. Finally, from our discussion of the coefficient of accommodation, it is clear that of those particles landing on a cold trap a fraction α must remain trapped. Hence, on any individual jump: probability of a particle being trapped = $K(1 - \alpha)\alpha$

where

α = probability of a particle escaping on a single jump.

Now since we assume that the jumps are uncorrelated, the above discussion must apply for all particles during a single jump. Hence $\alpha = 1 - \exp(-\bar{t}/\tau)$ where τ is a characteristic decay time for many-particle jumps. Therefore, the fraction, γ , of molecules that become permanently trapped on any jump compared to the total that either escape or are trapped on the jump is

$$\gamma = \frac{K\alpha(1 - \alpha)}{K\alpha(1 - \alpha) + \alpha} \quad (5)$$

However, the net number of particles in the

lunar atmosphere must remain constant for a constant escape probability α , and certainly does not change significantly during the time of a single jump. Hence, the evaporation rate must be exactly equal to the condensation rate plus the escape rate. Therefore,

$$\dot{E} = \dot{C} + (1 - \gamma)\dot{E}$$

or

$$\dot{C} = \gamma\dot{E} \quad (6)$$

and, from equations 1 and 6

$$\dot{m}_s = KA\dot{E}(1 - \gamma) = (1 - \gamma)\dot{m}_{s(\max)} \quad (7)$$

and from equation 5:

$$\dot{m}_s = KA\dot{E}\left(\frac{\alpha}{Ka(1 - \alpha) + \alpha}\right) \quad (8)$$

This relationship clearly demonstrates important physical restrictions on \dot{m}_s , the mass loss rate. For very large values of the decay time τ , the probability of escape α , and \dot{m}_s , are essentially zero. Thus, from equation 1, \dot{E} is equal to \dot{C} , and the partial pressure of a particular constituent in the lunar atmosphere is just the equilibrium vapor pressure corresponding to the temperature of the cold trap. For very small values of τ (rapid escape), α approaches 1, \dot{m}_s becomes $KA\dot{E} = \dot{m}_{s(\max)}$, and the partial pressure drops to zero. It is particularly important to note that the value $KA\dot{E} = \dot{m}_{s(\max)}$ is an absolute upper limit for the rate of removal of a condensed volatile from all the cold traps on the moon. As Figure 2 illustrates, \dot{m}_s is limited by Ka and by the evaporation rate of a particular volatile at 120°K, for example, rather than by the escape probability. For this reason, the evaporation rate data of Figure 1 are a much more meaningful guide to stability of volatiles on the lunar surface than are properties such as molecular weight, ionizability, etc., that enter into escape computations.

2.2. Transient loss of newly liberated volatiles during migration to cold traps. We must also consider the behavior of any new volatile material that might be liberated at an arbitrary point on the lunar surface either sporadically or steadily. Let us define

\dot{m}_f = mass influx of newly liberated volatiles for the entire lunar surface, in grams per second. This can be considered to be either a smooth function or a

series of random impulses of arbitrary amounts.

\dot{m}_{fl} = mass loss rate by escape during random migration of the molecules from source point to a cold trapping surface in grams per second.

If we assume again that the end point of molecular jumps are uncorrelated with beginning points, then the previous derivation for condensation rate applies to the present problem as well. In particular, the fraction of newly liberated molecules that get trapped on the first jump, or on any subsequent jump, is simply $\gamma = [Ka(1 - \alpha)]/[Ka(1 - \alpha) + \alpha]$. Hence

$$\dot{m}_{fl} = (1 - \gamma)\dot{m}_f$$

2.3. Basic equations describing mass loss. Since the mass rate of removal by evaporation from the cold traps is independent of the time of liberation of any new volatile materials from the lunar surface, the two loss rates \dot{m}_s and \dot{m}_{fl} are independent. Therefore, from equation 1 and 9:

$$\begin{aligned} \dot{m} &= \dot{m}_s + \dot{m}_{fl} \\ &= (1 - \gamma)(\dot{m}_{s(\max)} + \dot{m}_f) \end{aligned}$$

when \dot{m} is the total mass loss rate from the lunar surface in grams per second and ΔM is the accumulated loss over an interval of time $T_2 - T_1 = \Delta T$ is simply

$$\begin{aligned} \Delta M &= \int_{T_1}^{T_2} \dot{m} dt = (1 - \gamma) \\ &\quad \cdot (\dot{m}_{s(\max)} + \dot{m}_f) \Delta T \end{aligned}$$

The mass loss from cold traps only can be expressed as

$$\begin{aligned} \dot{m}_c &= \gamma\dot{m}_f - \dot{m}_s \\ &= \gamma\dot{m}_f - (1 - \gamma)\dot{m}_{s(\max)} \\ m_c(t) &= m_c(0) + \gamma t\dot{m}_f \\ &\quad - (1 - \gamma)t\dot{m}_{s(\max)} \end{aligned}$$

In the next section we will consider the calculation of the parameters of the equation, compute $m_c(t)$ and t comparable to the lifetime of the moon.

3. APPLICATIONS TO THE LUNAR SURFACE

3.1. Introduction. It is first necessary to state that the environment of the lunar surface

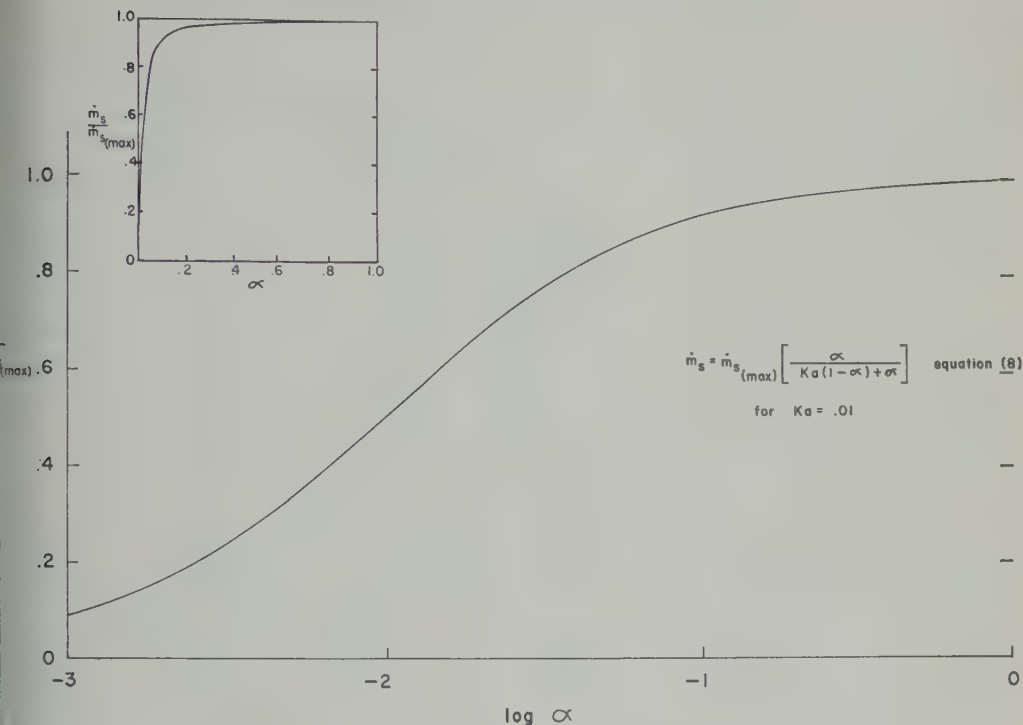


Fig. 2. Mass loss rate vs. escape probability.

ifies the two basic premises of our model. The existence of a rarefied atmosphere has been given observationally by Dollfus [1952], Elsmore and Whitfield [1955], and Öpik [1955].

The most recent studies by Dollfus [1956] of ionization and by Elsmore [1957] of radio occultations have verified the previous results and indicate that the lunar atmosphere is extremely rarefied. The maximum density consistent with the measurements of Elsmore yields a mean free path of the order of 10^4 km, which is sufficient to justify the use of ballistic trajectories and the mechanics of molecular transport in the lunar atmosphere, since the average jump length of molecules in equilibrium with the lunar surface is of the order of 10^2 km.

Infrared studies of the lunar surface by Hazel [1923], Pettit and Nicholson [1930], and Brown [1955] indicate that the shaded areas of the lunar surface have maximum temperatures of the order of 120°K . Thus 120°K represents an upper limit for permanently shaded areas; much lower temperatures would be expected if the surface temperature resulted principally from the balance on that surface between an

interior heat flow and radiative cooling. Simple computations indicate that the low conductivity of the dust layer as proposed by Jaeger [1953] and Wesselink [1948] will prevent even small areas of permanent shade from being heated up during the lunar day by horizontal conduction from adjacent illuminated areas.

The next two sections are devoted to a discussion of reasonable and limiting numerical values for K , the fraction of permanent shade, and α , the probability of escape. Using these values, the mass removal rates for common volatile substances will be compared, and we will conclude that water, in the form of ice, is the only common volatile that could be stable for a period of time comparable to age of the moon. We then discuss possible sources for the liberation of water from the lunar surface. Finally, a discussion is presented of the validity of the uncorrelated model as a description of migration on the actual lunar surface.

3.2. *The amount and distribution of permanently shaded areas on the lunar surface.* The obliquity of the moon's axis of rotation with respect to the ecliptic is only $1^\circ 32'$ [Allen, 1955]; there is

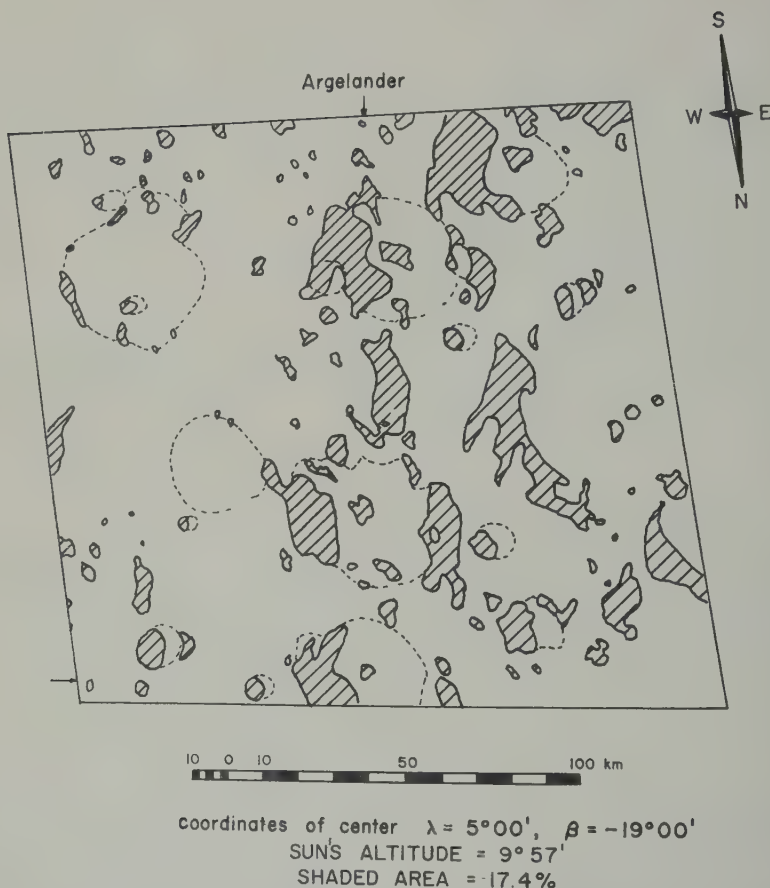


Fig. 3. Tracing of the shaded areas near the crater Argelander for the sun's altitude of $9^{\circ}57'$.

virtually no seasonal variation on the moon. Accordingly, many steep-walled craters at high lunar latitudes will be permanently shaded on the part of the wall and bottom of the crater nearest the equator. However, any terrestrial observation of the moon is necessarily made from a selenodetic direction that is also occupied by the sun at one time or another; hence, it is impossible to study the permanently shaded areas directly from the earth. One may attack the problem, however, in at least two indirect ways: (1) define some necessarily crude analytical model of a crater, compute its permanently shaded area as a function of latitude, and use some distribution function for lunar craters to arrive at an estimate of K , the fraction of permanently shaded craters for the whole lunar surface, or (2) use photographs of upland areas near the equator to obtain estimates of the

actual roughness in the polar areas, and, then, by measuring the variation in per cent shaded area as a function of the sun's elevation, estimate K . The first technique was applied by Watson, Murray, and Brown [1961] in order to get an absolute lower limit of $K = 10^{-4}$. The second approach, which is discussed briefly in the present paper and will be published elsewhere, yields an actual estimate of $K = 5 \times 10^{-4}$ (rather than just a lower limit).

Figure 3 is a tracing of the shaded areas in a small part of the lunar surface sufficiently near $\lambda = 0$, $\beta = 0$, so that any bias in projection is small. The character of the topography in this test area is quite similar to that in the polar areas: both display a dense array of craters of random-size distribution. Mountains, valleys, and other linear topographic features are almost entirely restricted to the borders of the main

TABLE 1. Estimation of Permanently Shaded Area on Lunar Surface

Latitude Zone, deg	Total Lunar Surface, %	Upland in Latitude Zone, %	Total Upland, %	Shaded Area at Noon, %	Permanently Shaded Area, % (for reduction factor 0.5)
90-80	1.53	100	1.53	27.5	0.21
80-70	4.54	100	4.54	8.5	0.19
70-60	7.41	94.5	7.00	2.2	0.08
60-50	10.04	78.5	7.88	1.0	0.03
50-40	12.36	77.5	9.57		
40-30	14.28	67.5	9.63		
30-20	15.76	57.0	8.99		
20-10	16.78	48.0	8.05		
10-0	17.30	47.0	8.13		
Total			65.3	Total	0.51%
				$K = 5 \times 10^{-3}$	

consequently, can be ignored when considering high latitudes where the permanently shaded areas are concentrated.

The relative proportion of upland as a function of latitude for the visible face of the moon is given in column four of Table 1. The percentage of the test area shaded at different elevation angles of the sun has been measured from other photographs of the same area at different lunar latitudes. Since the per cent shaded area of any crater on the equator at a sun's elevation of, for instance, 20° , is about equal to the shaded area that same crater would exhibit at local noon if it were at a latitude of $90^\circ - 20^\circ = 70^\circ$, it is possible to use the measurement of per cent shaded area vs. sun's elevation for an equatorial test area as a good estimate of the per cent shaded area at noon time of similar topography at other latitudes. The data from the photographs are plotted in the fifth column of Table 1 in this manner, except that the $1^\circ 32'$ annual variation in the sun's elevation has been taken into account. The annual variation, of course, is equivalent to increasing the sun's elevation at all latitudes by that amount. For any crater the maximum shaded area occurs at noon time, but the shadow then is still larger than the permanently shaded area. This is because the sun illuminates a part of the border of the area shaded by the noon-time shadow at other times during the lunar day, even though the elevation angle is less than at noon time. We take a value of K as an average reduction factor for lunar craters. The estimated permanently shaded area

vs. latitude as a percentage of the total lunar surface area is then merely the product of columns 4 and 5, reduced by the $\frac{1}{2}$ factor.

The main source of error leading to an overestimate of K would arise from a poor choice of the reduction factor; $K = 10^{-3}$ would seem to be an absolute lower limit. Similarly, the facts that (1) there are less maria on the reverse side of the moon than on the visible one *Barabashov, Michailov, and Lipski* [1961], and (2) that craters less than 1 km in diameter have effectively been ignored owing to limited resolution, might suggest that the true value of K is higher than 5×10^{-3} . An upper value of $K = 10^{-2}$ seems appropriate. Accordingly, we conclude that K lies between 0.1 and 1 per cent, most likely about 0.5 per cent.

3.3. *Escape mechanisms.* The most important escape mechanisms proposed so far are ionization, solar wind collision, and gravitation.

Öpik and Singer [1960] have calculated the half-life of krypton in an exosphere in contact with the lunar surface under solar ultraviolet radiation. They show that incident radiation will ionize the atmosphere and also raise the electrostatic potential on the lunar surface by photoionization of silicates. By calculating the screening length associated with this potential and comparing it with the characteristic 'scale height' of the atmosphere they calculated a half-life of 4.3×10^{10} secs ($\tau = 6 \times 10^{10}$ seconds). *Öpik and Singer* (private communication) have also estimated the decay time for photodissociation of water vapor in the lunar atmosphere by

TABLE 2. Escape Parameters

Volatile	μ	\bar{D} , km	\bar{t} , sec	Escape Probability, α		
				Solar Radiation	Solar Wind	Gravitational
Mercury	200	26	133	2×10^{-9}	1×10^{-3}	5×10^{-3}
Krypton	83	62	206		2×10^{-3}	2×10^{-3}
Sulfur dioxide	64	61	236		2×10^{-3}	3×10^{-3}
Carbon dioxide	44	118	285		3×10^{-3}	7×10^{-3}
Hydrogen chloride	36	142	315	4×10^{-3}	3×10^{-3}	4×10^{-3}
Water	18	287	446		4×10^{-3}	2×10^{-3}
Ammonia	17	304	458		5×10^{-3}	4×10^{-3}

solar radiation to be of the order of 10^5 seconds or less, but the relevant calculations have not yet been published.

Herring and Licht [1959] have estimated the reduction in a lunar atmosphere due to the influx of the high-energy protons in the solar wind. Since the protons are traveling with an assumed velocity of 10^8 cm/sec they have an incident energy of about 5 kev. Inasmuch as the energy necessary for escape of even the heavy volatiles from the lunar atmosphere is less than 5 ev it is apparent that the energy transferred by proton-molecule impact is more than sufficient to cause escape. The molecular diameter of the volatiles is 3×10^{-8} cm, hence the cross section for impact is about 10^{-15} cm². For a solar wind density of 10^2 particles/cm³, the decay time is $(10^{-15} \times 10^2 \times 10^8)^{-1} = 10^5$ seconds.

The least rapid rate of removal is by direct escape of that fraction attaining thermal velocities greater than the escape velocity. Spitzer [1952] has derived the equations for the escape time. We shall assume that the temperature at the base of the exosphere is the lunar surface temperature of 400°K. Then

$$\tau = \frac{\sqrt{6\pi}}{3g} \cdot V \frac{e^Y}{Y}$$

where

$$Y = \frac{3V_\infty^2}{2V^2}$$

and

- V_∞ = escape velocity.
- V = rms velocity associated with 400°K.
- For water, $\tau = 2.5 \times 10^8$ seconds.

Table 2 gives α , the escape probability, for

the mechanisms we have discussed, as well as the values for \bar{t} , the average jump time, and the average jump length for selected volatiles.

3.4. Mass loss rates of water and other volatiles. The mass loss rate per unit lunar surface area can be computed from equation 8. We shall use the value for K of 5×10^{-3} , and, assuming $a = 1$, we use Table 2 and the evaporation rates (Fig. 1) to compute \dot{m}_e . Our assumption that a is unity for ice is justified by the work of Tschudin [1946] who measured evaporation rates of ice in the range -60°C to -85°C and determined a value of .94 for a . The results are presented in Table 3 and illustrated graphically in Figure 4.

It is apparent that mechanical escape is too inefficient to be considered a significant loss mechanism. A comparison of the other loss mechanisms leads us to the conclusion that only water and mercury could remain trapped for an appreciable fraction of lunar history. Because of the extremely low abundance of free mercury on the earth's surface, it is reasonable to neglect it from further lunar consideration. Accordingly, water is the only volatile possibly retained in significant quantities on the lunar surface. We will use the mass loss rate associated with photodissociation as a reasonable estimate of the escape of water.

The total mass loss rates of water are computed for two values of α : a 'reasonable' one corresponding to $\tau = 10^5$ seconds, and a maximum one of $\tau = 0$ seconds (i.e., zero condensation) where values for the latter will be given in brackets following the values associated with the reasonable estimate.

In one billion years the mass loss rate from the cold traps is 4 g/cm² (9 g/cm²) (cm² refers to the whole lunar surface). This is equivalent

TABLE 3. Mass Loss Rate Per Unit Area of the Lunar Surface m_v/A (g/cm²/sec)

Volatile	Solar Radiation	Solar Wind	Mechanical	Maximum limited by evaporation
Mercury	1×10^{-7}	6×10^{-28}	2×10^{-96}	3×10^{-25}
Xenon		9×10^{-2}	1×10^{-28}	3×10^{-1}
Sulfur dioxide		1×10^{-9}	2×10^{-29}	4×10^{-9}
Carbon dioxide		1×10^{-6}	1×10^{-18}	1×10^{-5}
Hydrogen chloride	1×10^{-16}	1×10^{-4}	2×10^{-14}	3×10^{-4}
Water		1×10^{-16}	1×10^{-19}	3×10^{-16}
Ammonia		3×10^{-3}	5×10^{-11}	6×10^{-8}

removal of a total of 8 meters (18 meters) of ice from the cold traps. We now calculate the necessary amount of new water that would have been liberated from the whole lunar surface to replenish the loss from the cold traps. From equation 5 the mass addition rate to the cold traps owing to liberation of new volatiles is $\gamma \dot{m}_f$. Hence, for no net mass loss from the cold traps,

$$\gamma \dot{m}_f = \dot{m}_s \quad (14)$$

From equations 7 and 15

$$\dot{m}_f = \frac{1 - \gamma}{\gamma} \dot{m}_{s(\max)}$$

Using our 'reasonable' estimate for \dot{m}_s , the required liberation rate is 7 g/cm²/billion years.

3.5. Production of water on the lunar surface.

Actually all known silicate materials that fall on the earth from space appear to contain water in varying amounts, the greater part of which is probably chemically bound. The perthene, bronzite, and enstatite chondrites contain very low concentrations of water, starting at about 0.02 per cent, averaging about 0.25 per cent, and seldom exceeding 0.5 per cent. There is always the uncertainty that some of the observed water might be of terrestrial origin, but water does appear to be present in fresh samples which have been analyzed shortly after recovery. Concentrations range upward to as high as 20 per cent in some carbonaceous chondrites. In the latter cases there is little question that the water is extraterrestrial. These meteorites represent some 3 per cent of all observed meteorites.

The visible water now on the surface of the moon, which appears to have been present originally in chemically bound form, amounts to somewhat over 280 kg/cm² of the earth [Rankama

and Sahama, 1950]. Approximately 6 kg of water per square centimeter of the earth appears to be bound in the sediments and, judging from the observed water contents of basaltic materials, an additional 50 kg/cm² may still reside in the crust above the Mohorovicic discontinuity. Some 30 kg of water per square centimeter has perhaps been lost over geologic time as the result of dissociation in the atmosphere followed by the escape of hydrogen. Thus, if the earth is well-degassed below the Mohorovicic discontinuity, the primitive earth would have contained some 370 kg of water per square centimeter. On this basis the original earth material would have contained about 0.03 per cent water, an amount comparable to the apparent abundance of bound water in many chondrites of the noncarbonaceous variety.

Had primitive lunar material contained an equivalent concentration of water, the moon would have contained originally an amount of water equal to about 60 kg of water per square centimeter of the lunar surface. This amount of water is equal to about 10⁴ times the amount of water we estimate to have been lost per billion years from the lunar surface. Accordingly, it seems reasonable to assert that there should still be detectable amounts of ice in the permanently shaded areas of the moon if the moon has undergone a bulk chemical differentiation as small as about one one-thousandth that of the earth. We may conclude, then, that ice should prove to be a most sensitive mineral indicator of the degree of chemical differentiation of the moon, a far more sensitive guide than the percentage of silica in the surface rocks, for instance.

It should be pointed out that accretions of meteoritic water upon the moon are probably small when compared with the amount of water liberated from the interior of the moon itself.

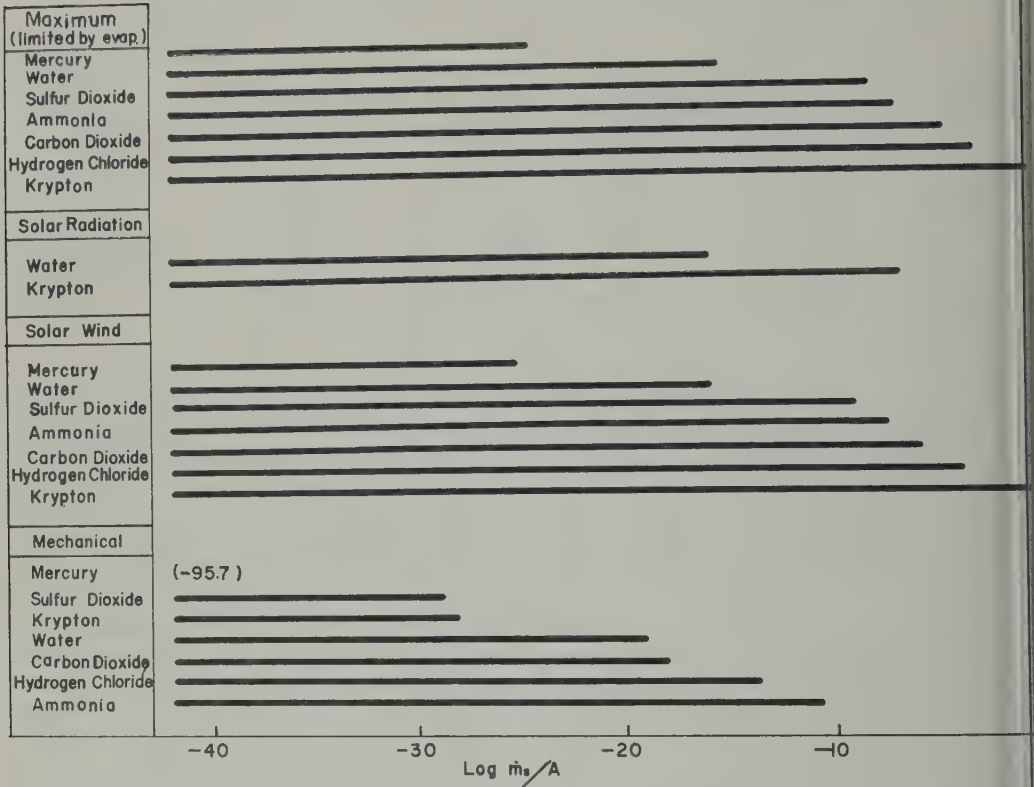


Fig. 4. Mass loss rate in $\text{g/cm}^2/\text{sec}$ for various escape mechanisms.

The frequency distribution curve for meteoritic masses indicates that at the present rate of influx perhaps 20 grams of meteoritic matter may have been deposited upon every square centimeter of the lunar surface during geologic time [Brown 1960, 1961]. This could have resulted in the liberation of no more than about 0.2 grams of water per square centimeter of the moon, largely from the carbonaceous chondrites.

It should also be pointed out that the flux of 'meteorites' composed primarily of ice of cometary origin is unknown. Such objects might well be fairly abundant, but their very nature would preclude our observing them reliably from the earth's surface.

Finally, it is interesting to compare the required liberation rate of $7 \text{ g/cm}^2/\text{billion years}$ computed in the previous section with the thickness of a lunar crust containing 0.03 per cent water which would have to be differentiated to produce that much water. The required thickness is 78 meters, only about one-half of 1 per

cent of the average crustal thickness on the earth.

3.6. *Validity of migration model.* It is not possible to return to the question of the relative distance between cold traps as compared to the average distance covered by a migrating molecule in a single jump. It will be recalled from section 3.5 that our theory assumes that the beginning and end points of a single jump are uncorrelated—particular, that a molecule will not find itself in an area devoid of cold traps which is larger compared to its average jump length. Actually, such a case exists for water molecules that are released in an equatorial area, but we will not show that this departure from our model is not significant. The 'barren' equatorial areas are approximately compensated by the polar areas where the proportion of cold traps is much larger than indicated by our average value of 1 for the entire moon, and the jump distances are less because of the lower average surface temperature.

The opposite extreme of our uncorrelated

lel would be the case where the jump lengths are so short compared with the separation between cold traps that there is little migration, that the beginning and end points are highly correlated in latitude. For instance, very few molecules released in the equatorial areas could reach the higher latitudes to be trapped. However, molecules released in the polar areas could be confined to these higher latitudes where larger K would result in a higher proportion being captured. Hence, for the correlated case the loss rate from the cold traps is reduced about 50 per cent (see Table 1). We shall examine the trapping of newly liberated volatiles for the correlated model by considering two cases: uniform liberation over the lunar surface, and extreme case of liberation from the equatorial areas only.

The uniform liberation results in the loss rate being migration to the cold traps being increased about 50 per cent. The net effect when combined with the lower loss rate from the cold traps is to increase the required liberated flux to balance the cold-trap loss by about 10 per cent on the average. Clearly this difference is insignificant in terms of the uncertainties in K and α , and the use of the uncorrelated model is sufficiently accurate for our purposes.

The correlated and uncorrelated models do not differ significantly, however, in regard to preferential loss of volatiles liberated in the equatorial regions. A rigorous treatment of this problem is not solved, but the following analysis is sufficient to justify the uncorrelated model.

From Table 1 it will be seen that the fraction of permanently shaded area becomes insignificant at latitudes less than about 50°. Accordingly, the worst case for a migrating molecule to reach a cold trap is its release at the equator, 1500 km from the 50° latitude zone where the truly uncorrelated model begins to hold. The migration is resumed to be a random walk in arbitrary directions with an average jump distance, for a single molecule, of 280 km (Table 2). The average component of northerly or southerly migration per jump is then $.707 \bar{D} = 200$ km. We wish to find the expected number of jumps, N , for the molecule to reach either 50°N or 50°S. This condition can be expressed as

$$\Delta N = \frac{1500}{200} = 7.5 \quad (16)$$

where ΔN is the difference between the number of northerly component and the number of southerly component jumps which comprise N . But since the direction of jump is presumed to be random

$$\Delta N = N^{1/2} \quad (17)$$

and

$$N = (7.5)^2 = 56 \quad (18)$$

Since on each jump the probability of escape is α , the probability of reaching 50° latitude by random walk without escaping is just

$$(1 - \alpha)^N = 1 - N\alpha \quad (19)$$

since $N^2\alpha^2$ is negligible.

From Table 2, $\alpha = 4 \times 10^{-3}$

$$\therefore (1 - \alpha)^N = 0.78 \quad (20)$$

Hence, for the most extreme case, 22 per cent m_r at the equator would be lost before entering into the uncorrelated zone. The loss decreases for sources located between 0° and 50° latitude, and it is clear that the net preferential selection effect is insignificant. Thus, the uncorrelated model is indeed a valid description of migration as well as total trapping of volatiles on the lunar surface.

We have deferred until the present section a discussion of the effect on the mass loss rate of variations in the parameters K , a , and α . It is clear from equations 8 and 9 that K and a only occur as the product Ka . If we double the value of Ka , we increase γ from .555 to .714. Thus the total mass loss from the cold traps is increased but the fraction of liberated volatiles lost from the atmosphere is decreased; the net effect is to reduce the required liberated flux to balance the cold trap loss by a factor 2. If the value of Ka is halved, or α is doubled, then γ decreases from .555 to .385, and the net effect is to increase the required liberated flux by a factor 2. Obviously these changes in mass loss are insignificant compared to changes in the evaporation rates with temperature. We have used 120°K for the cold trap temperature, with the understanding that it represents a maximum. If a more likely temperature of 100°K is used, the mass loss rate of water from the cold traps is reduced by four orders of magnitude, and this loss is insignificant over the lunar lifetime

(equivalent to a maximum loss from the cold trap surfaces of 2-mm thickness per billion years).

Finally, we can summarize our conclusions from this study by stating that the permanently shaded areas are such an efficient trap that the discovery of no ice in these areas on the moon implies the stringent limitations of an extremely low production of water from the lunar surface during its lifetime.

The atmospheric density that corresponds to the vapor pressure of water at 120°K is 3.5×10^4 molecules/cm³. It is interesting to note at this point that the degree of ionization of water produced by the solar wind is approximately the escape probability $\alpha = 4 \times 10^{-3}$, since this represents the fraction of molecules that are ionized and escape; the ion density would be 10^2 /cm³. Clearly this result is compatible with the density of 10^3 electrons/cm³ observed by *Elsmore* [1957]. In addition, the logarithmic variation of vapor pressure with inverse temperature implies that the vapor pressure decreases very rapidly for a small temperature decrease. Since 120°K is an upper limit, we should actually expect to find even lower densities.

It has been suggested in personal communications to the authors that sputtering [see *Wehner*, 1955] might be a significant erosional mechanism acting on the ice in the cold traps. Generally the threshold energy for sputtering is at least an order of magnitude greater than the latent heat of vaporization. Recent calculations by *Brandt* [1961] suggest that *Chamberlain's* [1961] model of the solar corona with velocities of a few kilometers per second and a density of around 30 electrons/cm³ at the earth's orbit is in agreement with the deflection of comet tails. The resulting energy flux is many orders of magnitude lower than the radiative heat transfer from the cold trap surfaces and seems to exclude the possibility of sputtering having an appreciable effect on the mass removal rate of ice from the cold traps.

4. SUMMARY

We have developed a model of the stability of volatiles on the lunar surface. This model is based upon the presumed existence of permanently shaded areas that act as cold traps for volatiles and a rarefied atmosphere that

permits the use of simple dynamic trajectories over the lunar surface. Both these assumptions are validated by observations of the surface. The model yields simple expressions for the mass removal rates of volatiles from the traps and their migration to the cold traps following release from the lunar crust, particularly:

$$\dot{m} = (1 - \gamma)(AK\dot{E} + \dot{m}_f)$$

where

\dot{m} = total mass removal rate of a volatile from the moon (g/sec).

A = surface area of moon (cm²).

K = fraction of surface in permanent shadow.

\dot{E} = evaporation rate of volatile at temperature of cold trap (assumed to be 120°K in calculation) (g/cm²/sec).

\dot{m}_f = mass influx rate for liberation of volatile material on lunar surface (g/sec).

γ = $Ka(1 - \alpha)/[Ka(1 - \alpha) + \alpha]$ = ratio of trapped fraction to trapped plus escaping fractions.

where

a = the accommodation coefficient of volatiles on cold traps.

α = escape probability on a single jump.

The parameters of the model have been examined in order to estimate reasonable and also limiting values for water on the lunar surface. These are:

K = 5×10^{-3} with a maximum range of 1×10^{-3} to 1×10^{-2} .

a = approximately 1.

α = 4×10^{-3} for photodissociation of water.

α_{\max} = 1 for maximum loss by evaporation.

The application of the model to cometary volatiles leads us to the conclusion that water is relatively stable over periods comparable to the lunar lifetime. We have calculated the total mass loss rate of water from the cold traps to be approximately 4 g/cm²/billion years for the reasonable estimate. We also estimated the average liberation rate of water over the lunar surface necessary to condense in the cold traps and balance the loss to be 7 g/cm²/billion years.

Finally we have discussed the validity of our model and the effects of possible variations in parameters α , K , and a . The model was found to be quite adequate, since the variation of parameters produced effects which are negligible compared to the change in mass loss rate produced by reducing the cold trap temperatures from the maximum of 120°K to a more reasonable temperature of 100°K.

These results imply that the lack of ice in the cold traps today would require that the liberation of water from the moon during its lifetime must have been extremely low compared with that on the earth and hence the implication that chemical differentiation to form a lunar crust must have been negligible.

Acknowledgement. This work has been supported in full by the National Aeronautics and Space Administration under Grant NsG 56-60.

REFERENCES

- Allen, C. W., *Astrophysical Quantities*, Athlone Press, p. 162, 1955.
- Abashov, N. P., A. A. Michailov, V. N. Lipski, Editors of *Atlas of the Opposite Side of the Moon*, Academy of Sciences, USSR, 1961.
- Adams, J. C., On the study of comet tails and models of the interplanetary medium, *Astrophys. J.*, **133**, 1091-1092, 1961.
- Altmann, P. W., High pressure effects, in *American Institute of Physics Handbook*, 1957.
- Brown, Harrison, The density and mass distribution of meteoritic bodies in the neighborhood of the earth's orbit, *J. Geophys. Research*, **65**, 1679-1683, 1960.
- Brown, Harrison, Addendum: The density and mass distribution of meteoritic bodies in the neighborhood of the earth's orbit, *J. Geophys. Research*, **66**, 1316-1317, 1961.
- Chamberlain, J. W., Interplanetary gas, 3, A hydrodynamic model of the corona, *Astrophys. J.*, **133**, 675-687, 1961.
- Clusafus, A., Nouvelle recherche d'une atmosphere au voisinage de la lune, *Comptes rendus, Academie des Sciences*, **234**, 2046-2049, Paris, 1952.
- Clusafus, A., Recherche d'une atmosphere autour de la lune, *Ann. d'Astrophys.*, **19**, 71, 1956.
- Croft, B., and G. R. Whitfield, Lunar occultation of a radio star and the derivation of an upper limit for the density of the lunar atmosphere, *Nature*, **176**, 457-458, 1955.
- Croft, B., Radio observations of the lunar atmosphere, *Phil. Mag.*, **2**, 1040-1046, 1957.
- Fermi, L., Gases at low densities, in *Thermodynamics and Physics of Matter*, Vol. 1, *High Speed Aerodynamics and Jet Propulsion*, edited by F. D. Rossini, Princeton University Press, pp. 742-744, 1955.
- Herring, J. R., and A. L. Licht, Effect of solar wind on the lunar atmosphere, *Science*, **130**, 266, 1959.
- International Critical Tables*, **3**, 201-246, McGraw-Hill Book Company, New York, 1928.
- Jaeger, J. C., The surface temperature of the moon, *Australian J. Phys.*, **6**, 10-21, 1953.
- Knudsen, M., Die molekulare Wärmeleitung der Gase und der akkommodationskoeffizient, *Ann. Physik*, **34**, 593-656, 1910.
- Kuiper, G. P., Planetary atmospheres and their origin, in *The Atmosphere of the Earth and Planets*, revised edition, edited by G. P. Kuiper, University of Chicago Press, Chicago, p. 367, 1952.
- Langmuir, I., The condensation and evaporation of gas molecules, *Phys. Rev.*, **3**, 141-147, 1917.
- Loeb, L. B., *Kinetic Theory of Gases*, 1st ed., McGraw-Hill Book Company, Inc., New York, pp. 240-337, 1927.
- Menzel, D. H., Water-cell transmissions and planetary temperatures, *Astrophys. J.*, **58**, 65-74, 1923.
- Öpik, E. J., The density of the lunar atmosphere, *Irish Astron. J.*, **3**, 137-143, 1955.
- Öpik, E. J., and S. F. Singer, Escape of gases from the moon, *J. Geophys. Research*, **65**, 3065-3070, 1960.
- Pettit, E., Radiation measurements on the eclipsed moon, *Astrophys. J.*, **91**, 408-420, 1940.
- Pettit, E., and S. B. Nicholson, Lunar radiation and temperatures, *Astrophys. J.*, **71**, 102-135, 1930.
- Rankama, K., and Th. G. Sahama, *Geochemistry*, University of Chicago Press, Chicago, 1950.
- Sinton, Wm., Observations of solar and lunar radiation at 1.5 millimeters, *J. Opt. Soc. Am.*, **45**, 975-979, 1955.
- Spitzer, L., Jr., The terrestrial atmosphere above 300 km, in *The Atmosphere of the Earth and Planets*, revised edition edited by G. P. Kuiper, University of Chicago Press, Chicago, pp. 239-244, 1952.
- Tschudin, K., Die Verdampfungsgeschwindigkeit von Eis, *Helvetica Physica Acta*, **19**, 91-102, 1946.
- Urey, H. C., *The Planets, Their Origin and Development*, Yale University Press, New Haven, pp. 17-18, 1952.
- Vestine, E. H., Evolution and nature of the lunar atmosphere, *Rand Research Memo. RM-2106*, 1958.
- Watson, K., B. Murray, and Harrison Brown, On the possible presence of ice on the moon, *J. Geophys. Research*, **66**, 1598-1600, 1961.
- Wehner, G. K., Sputtering by ion bombardment, in *Advances in Electronics and Electron Physics*, **7**, 239-298, Academic Press, 1955.
- Wesselink, A. J., Heat conductivity and the nature of the lunar surface material, *Bull. Astr. Inst. Netherlands*, **10**, 351-363, 1948.

(Manuscript received June 9, 1961;
revised June 30, 1961.)



Geomagnetic and Solar Data

J. VIRGINIA LINCOLN, EDITOR

Central Radio Propagation Laboratory
National Bureau of Standards
Boulder, Colorado

INTERNATIONAL DATA ON MAGNETIC DISTURBANCES

This report continues the series which has appeared regularly in this JOURNAL since Volume 65, No. 3, 295 (1949). Please refer to that first report for an explanation of the data given, and Volume 59, No. 3, 423 (1954) for the definition of Ap.

Note: Additional and final Geomagnetic and Solar Data appears in due course in the following international publications: *Quarterly Bulletin on Solar Activity*, International Astronomical Union, c/o Eidgen. Steinwarte, Zurich, Switzerland; *IAGA Bulletins, Geomagnetic Indices K_p and C_p*, by J. Bartels, A. Romaña and J. Veldkamp, International Union of Geodesy and Geophysics, Association of Geomagnetism and Aeronomy, c/o V. Laursen, Meteorologisk Institut, Charlottenlund, Denmark.

SUDDEN COMMENCEMENTS AND SOLAR-FLARE EFFECTS, FOURTH QUARTER, 1960

Preliminary Report of Sudden Commencements
S.c.'s given by ten or more stations are indicated. Times are mean values obtained from multiple magnetograms.

Sudden commencements followed by a magnetic storm or a period of storminess (s.s.c.)

1960 October 04d 11h 32m: Es Hl St (si: Fu; sfe: Db).—04d 14h 27m: Tr Db Lg Ci Qu MB Va.—06d 02h 37m: twenty-one (si: 18; si: 3).—24d 14h 52m: fifty-four.—05h 44m: Vi Te.

1960 November 03d 22h 28m: Ag Fr.—07h 18m: thirty-four (ssc: 21; si: 12; sfe: 11).—11d 00h 34m: twenty-three (ssc: 21; si: 10; si: 1; bs: 1).—12d 13h 49m: forty-nine (ssc: 44; si: 4; sfe: Va?).—12d 18h 45m: Gu AA (si: Mb Ka Ky).—15d 13h 04m: forty-four

(ssc: 42; si: 2).—15d 22h 00m: Fu Qu Ta Hr (si: Ci SF).—21d 06h 31m: thirty-three (ssc: 32; si: 1).—24d 12h 32m: Cm Fr Qu MB? Mc Hn Va (si: Bi).—30d 19h 09m: fifty (ssc: 49; si: 1).—30d 23h 58m: Cm Ci Fr Qu Ta AA Gn Hr Am (si: Mb Ka Ky Ho).

1960 December 07d 18h 04m: forty-five (ssc: 42; si: 2; b: 1).—14d 09h 13m: Ta Mc Hn (si: Lg).—15d 14h 13m: Hl Wn Sw V1 Pr (bp: Ta).—18d 05h 14m: Tr Db Ks Gu AA Hn PM (si: Fu Ta Mc Am).—25d 20h 02m: forty-three (ssc: 27; si: 16).—27d 05h 10m: Ka Qu Tn.

Sudden impulses found in the magnetograms (s.i.)

1960 October 02d 11h 56m: Nu Hl Sw Fu (bp: Ho).—04d 19h 06m: Ka Ky Qu.—06d 16h 52m: Mb Ka MB.—07d 01h 01m: Fr MB (bp: Hr).—07d 08h 44m: Wn V1 Pr Fu Ky AA (ssc: Wi Ta; b: To; bp: Hn).—07d 14h 44m: Es St Wn Pr Qu (bs: V1).—13d 21h 47m: forty-two (si: 23; ssc: 19).—23d 17h 25m: Cm Lg MB (bs: Qu; sfe: Tl).—25d 15h 50m: Es V1 Fu SF (ssc: Le St).—25d 20h 07m: Wn Pr (bp: Hr; pt: Cm).

1960 November 11d 02h 12m: Ta Mc Hr.—11d 03h 44m: Mb Ka Ky Ta Mc.—13d 06h 35m: Co Fr.—13d 10h 22m: Es Wn Mb Fr Ka Ky (ssc: Fu).—14d 18h 59m: Le Es (bs: V1).—19d 10h 20m: AA (ssc: Mc).—19d 12h 04m: Cm Aq.—21d 18h 22m: Nu Es St Wn V1 Pr Fu (bp: Eb; pt: Ta).—21d 21h 47m: fifteen (si: 13; ssc: 2).—22d 11h 00m: Hl Sw Db (sfe: Ci).—24d 20h 52m: Fu Lg Tl SF Ka Ta Ho PM (ssc: Vi Fr Tu Gu Am).

1960 December 01d 01h 18m: Ka Ky.—01d 02h 56m: Es Cm Ta.—05d 19h 43m: St Ha Lg Tl Qu (ssc: Cm Db Aq Va; bs: Mc; pt: Fu).—07d 22h 23m: Tu Qu Ta To.—08d 04h 52m: Ta AA Va (ssc: Qu).—14d 10h 57m: fifteen (si: 10; ssc: 5).—14d 15h 02m: thirteen (si: 11; ssc: 2).—16d 09h 20m: Si Vi (bps: Am).—17d 23h 27m: Sw Be Lg Tl Ta Mc El (b: Tn?; bps: V1; pg: Bi; sfe: Ho Gu PM).—20d 08h 34m:

ten.—26d 15h 57m: El Hr (ssc: MB).—27d 15h 16d 18h 19m: Hu.—29d 10h 24m—12h 14m:
15m: Wn Pr Fu SF Hr (ssc: Le; bp: Vl El; 31d 11h 20m: Fr.

Preliminary report on Solar-flare effects (s.f.e.)

Effects confirmed by ionospheric or solar observations are in italics.

1960 October 02d 14h 06m: Hl.—05d 20h 44m: Te.—11d 11h 12m: Db Eb? Ci Ta Gu Mc? Hu (si: Hl Tl).—11d 14h 06m: Te (si: Tl; bp: Fu).—14d 10h 15m—10h 58m: Ks Gu Mc.—14d 18h 39m: Hu.—15d 11h 14m—11h 24m: Es.—

1960 November 08d 06h 41m: Qu (si: Hl).—08d 11h 04m: Le Db Mc? (b: Bi).—08d 17h 10d 13h 21m: Vl.—10d 10h 17m: Tn Hu.—10d 15h 10m: Hu.—12d 13h 25m: Es Wn. Cm Pr Eb Fr Hr (ssc: 10; si: 1; bs: 1).—15d 21m: Mb Ka Ky Gu Hn Gn To (ssc: PM Ar).—22d 17h 22m: Tu (si: Vi).

1960 December 01d 09h 40m: Ta.—06d 54m—12h 04m: Tn (ssc: Ci; si: Sw).—19d 57m—13h 07m: Va.—24d 16h 19m: Ho (pt:

SELECTED GEOMAGNETIC AND SOLAR DATA

Kp, Ci, Cp, Ap, K_F, Rz and Selected Days
May 1961

Day ¹	Three-hour Range Indices <i>Kp</i> ²								Sum	Prel. ³ <i>Ci</i>	<i>Cp</i> ⁴	<i>Ap</i> ⁵	3-hr. Range Indices <i>K_F</i> ⁶		Prov. ⁷ <i>Rz</i>		
	1	2	3	4	5	6	7	8					Values	Sum			
Q	3+ 4- 1+ 3+ 3+ 3o 1- 3+	6- 5- 3- 3o 1o 2- 1- 1-	0+ 1+ 1+ 1o 1- 1- 1o 0+	0o 0+ 1o 1o 1- 2- 3o 4+	2+ 3- 4- 2+ 2+ 5+ 4- 4-	22o	20o	7-	12o	26o	0.9	0.9	0.8	15	4323 3213	21	102
											0.0	0.1	1.0	19	5522 1110	17	84
											0.6	0.4		4	0222 0111	09	72
											1.1	1.0		8	0011 0135	11	55
														20	3343 2434	26	42
D	3- 5o 5o 5- 5- 2o 4- 4-	4- 3o 5- 4- 3o 4- 2+ 3+	2+ 2+ 3- 2o 2o 2+ 4- 3o	1+ 3- 4o 5- 3- 3o 1+ 2o	2+ 1+ 2- 2o 2- 1o 2- 3+	31+	27+	20+	22-	15o	1.2	0.9	1.3	30	2443 4333	26	36
											1.0	1.0	0.6	20	4343 3333	26	31
											0.7	0.6		11	2332 2233	20	28
											0.8	0.9		15	1345 2223	22	44
											0.5	0.4		8	3122 1124	16	46
Q	4- 3o 3+ 3- 4- 4- 4+ 3-	4- 4- 3o 3- 2o 3o 3- 3+	5+ 5- 3o 3o 4- 2+ 2+ 2o	3o 4- 3- 3o 2- 1+ 1- 1-	0o 0+ 0+ 1+ 1- 1+ 2- 2+	27o	24o	26+	17-	8o	1.0	0.8	1.0	19	3333 3343	25	45
											0.8	0.9		15	4433 2233	24	56
											1.1	1.1		22	5533 3322	26	52
											0.5	0.6		10	3323 1111	15	46
											0.2	0.1		4	0001 1013	06	38
D	4o 5+ 5- 3o 4+ 4o 3o 3-	3o 2- 1+ 2+ 3- 2+ 0+ 1o	1o 1- 1o 1o 1- 0+ 0+ 0+	2- 2o 2- 2o 2o 2o 2o 3+	3+ 4+ 4- 2o 2+ 2o 2+ 1-	31o	15-	5+	17-	21-	1.3	0.5	1.2	28	3443 4333	27	31
											0.5	0.4		8	3212 2212	15	23
											0.0	0.1		3	1111 0001	05	44
											0.6	0.4		8	2323 2233	20	47
											0.9	0.8		13	3442 2121	19	59
Q	1+ 1o 1+ 1o 1o 1o 2+ 2o	3- 2+ 2+ 2- 2- 3o 4- 2o	2+ 4- 2+ 3- 4- 4- 3- 1+	3o 2+ 2- 1+ 2- 2o 1o 2+	4- 6- 6- 4- 4+ 4- 4+ 2o	11o	19+	22+	15+	33o	0.3	0.7	0.2	5	1111 0123	10	58
											0.6	0.6		11	3322 2333	21	66
											0.9	0.8		14	3422 3332	22	74
											0.5	0.4		8	3332 1123	18	78
											1.3	1.3		34	3553 3332	27	72
Q	4- 1o 2o 1+ 2+ 1o 1- 1+	2+ 2+ 2+ 1o 1+ 1o 0+ 1o	1o 1- 1o 2- 3o 3o 3+ 3o	1o 3- 2- 0+ 0+ 1- 1+ 1-	1+ 1+ 2+ 3- 3- 3- 1+ 2+	13+	12-	17o	9-	17-	0.3	0.3	0.4	7	4120 2221	14	47
											0.8	0.6		6	2331 0111	12	41
											0.1	0.2		10	2112 3234	18	38
											0.6	0.5		5	2320 1111	11	36
														9	1233 2223	18	41
D	4- 3+ 4- 4- 4o 4o 3+ 4-									29+	1.1	1.1	22	4443 3433	28	24	
Means:										0.69	0.66	13			50.2		
No. of days:										31	31	31			31		

s:

Five quiet days (Q), ten quiet days (Q or q), five disturbed days (D) selected by Committee on Characterization of Magnetic Disturbances, J. Veldkamp, Kon. Nederlandsch Meteorologisch Institut, DeBilt, and.

Geomagnetic planetary three-hour-range indices *Kp* prepared by Committee on Characterization of Magnetic Disturbances, J. Bartels, Chairman, University, Göttingen, Germany.

Preliminary magnetic character-figures, *Ci*, prepared by J. Veldkamp.

Magnetic character-figures, *Cp*, prepared by J. Bartels.

Average amplitudes *Ap* (unit 2γ), prepared by J. Bartels.

Fredericksburg three-hour-range indices *K* ($K_9 = 500\gamma$); scale-values of variometers in γ/mm : $D = 2.7$; $Z = 2.5$, $Z = 3.3$) prepared by Robert E. Gebhardt, Observer-in-Charge, Fredericksburg Magnetic Observatory, Corbin, Virginia.

Provisional sunspot-numbers (dependent on observations at Zurich Observatory and its stations at Arno and Arosa) prepared by M. Waldmeier, Swiss Federal Observatory, Zurich, Switzerland.

Letters to the Editor

Radiation Measurements in the Slot between the Van Allen Belts to an Altitude of 1415 Kilometers

C. A. GÜRTLER

*NASA Langley Research Center
Langley Field, Virginia*

Satellite experiments on the radiation belts surrounding the earth have indicated the flux of high-energy particles decreases substantially in the slot between the inner and outer Van Allen belts. Relatively few independent measurements have been made in this region.

During the development of the Scout vehicles by NASA Langley Research Center, a number of performance tests were scheduled for firing from Wallops Island, Virginia. The trajectories chosen provided a means of obtaining radiation data in the slot area. A radiation sensor was placed in the payload of the Scout ST-1, which was fired July 1, 1960.

The radiation measuring instrument consisted of an Anton-type 114/6993 Geiger-Muller tube as the detector, and a circuit chassis which consisted of the supply batteries, a transistorized pulse-shaping network, a detecting circuit, and a transistorized high-voltage power supply. The d-c output of this chassis was fed into a subcarrier oscillator with a center frequency of 1.3 kc (IRIG Band 5). Data were transmitted continuously by means of an FM FM telemeter. Full-scale output voltage was obtained with an input of 30,000 counts/min. The instrument was designed to have a nonlinear output to obtain greater sensitivity in the lower counting rates.

The Anton-type 114/6993 Geiger-Muller tube responds to gamma radiation and has an efficiency of 85 per cent for fast charged particles. The effective cathode dimensions of the Geiger tube are 0.6 inch in diameter by 1.125 inches long.

The radiation sensing system was calibrated with a cobalt 60 source to determine the response to gamma radiation. Figure 1 shows a plot of milliroentgens per hour as a function of subcarrier oscillator frequency. A calibration fixture was fabricated that allowed checking this

calibration after the instrument had been installed in the vehicle. A final calibration of the radiation sensing system was made hours before firing by applying the flux sensor on the launching tower and receiving information at the ground station.

The response of the system to particle was obtained by substituting a pulse generator in place of the Geiger tube. Electrical inputs were fed into the input of the system and output frequency of the subcarrier oscillator was counted. Full-scale output of the subcarrier oscillator was obtained with an input of 30,000 counts/sec. With the relatively low full-scale range the counter had no appreciable dead

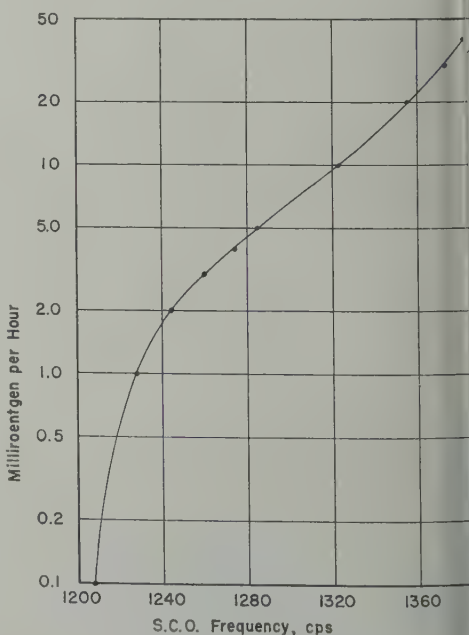


Fig. 1. Variation of milliroentgens per hour as a function of subcarrier oscillator frequency.

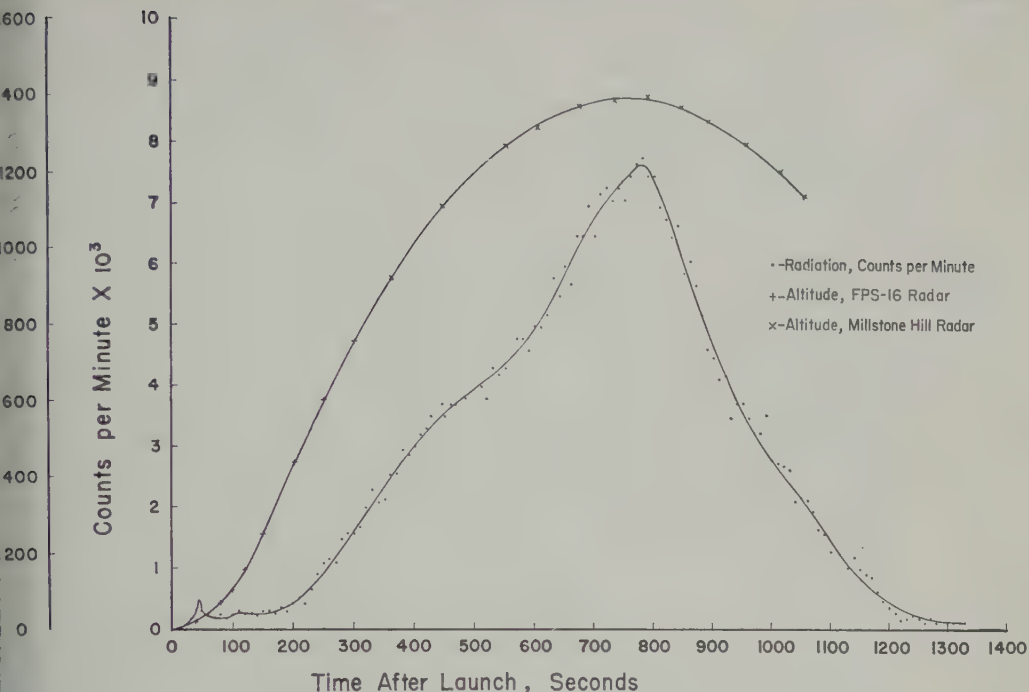


Fig. 2. Altitude in kilometers and radiation counts per minute as a function of time after launch in seconds.

ts. The calibration curve in Figure 1 may be converted to read counts per second as a function of subcarrier oscillator frequency by multiplying the milliroentgens per hour by 10. The radiation sensor was placed in the vehicle at the fiberglass nose of the payload chamber. The thickness of the fiberglass in the nose of the payload chamber was $\frac{1}{2}$ inch which would provide a shield of 2.22 g/cm^2 . The Geiger-tube window has a mass of 30 mg/cm^2 , so the total shield looking forward through approximately 1 steradian was 2.25 g/cm^2 . This would permit the counting of electrons with energies greater than 4.3 Mev, and protons with energies greater than 48 Mev. The cylindrical section of the payload container was fabricated of fiberglass 0.15 inch thick. This section had a mass of 0.6 g/cm^2 and would have permitted the counting of electrons with energies greater than 4 Mev and protons with energies greater than 48 Mev. Most of the area behind the plastic shield was filled with instrumentation (mass greater than 20 g/cm^2) so the look-angle through the cylindrical section of the payload container was approximately 0.5π steradian.

Results from the flight data received by the radar and telemetry have been plotted in Figures 2, 3, and 4.

Figure 2 is a plot of altitude in kilometers and counts per minute as a function of time after launch in seconds.

To obtain measured altitude from launch through the maximum altitude, it was necessary to use data from both the FPS-16 radar located at Wallops Island and the MIT Millstone Hill radar. The FPS-16 radar made measurements from launch to 611 seconds after launch at which time the signal was lost. The Millstone Hill radar picked up the vehicle at 220 seconds after launch and tracked it through maximum altitude. The altitude used in this plot was measured by the FPS-16 radar for the first 300 seconds after launch and the Millstone Hill radar from 300 seconds to 1050 seconds.

Radiation data was recorded from launch until the telemeter signal was lost just before splash (1326 seconds). The radiation count varied from normal background (20 to 30 counts/min) at launch to a maximum of 7,800 counts/min which occurred at the maximum

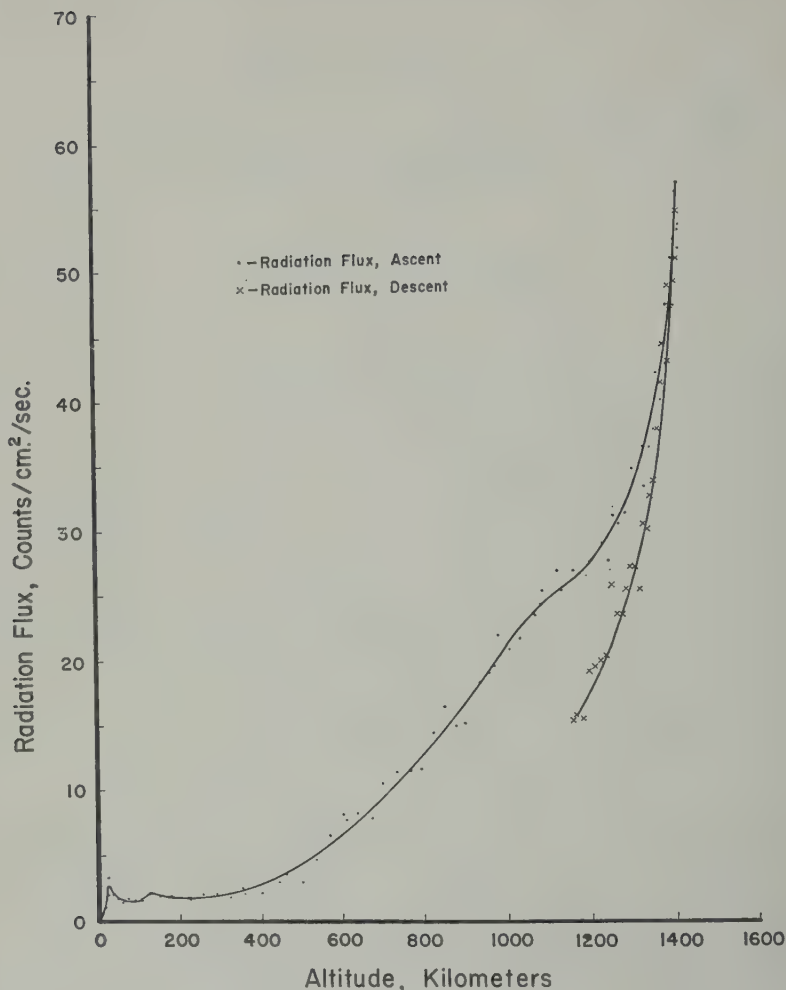


Fig. 3. Radiation flux as a function of altitude in kilometers.

altitude of 1415 km. A substantial portion of the count rate observed may be due to the bremsstrahlung of relatively low energy electrons.

During ascent the radiation count rate increased with altitude up to 18 km, at which time a peak rate of 450 counts/min was reached. The count rate then decreased to an average rate of 250 counts/min up to an altitude of 375 km. From 375 km altitude to 1100 km altitude the count rate increased at a fairly uniform rate. There appears to be a decrease in rate between 1100 and 1250 km, and then the rate increases rapidly reaching a maximum counting rate of 7800 counts/min at the maximum altitude of 1415 km. The count rate

doubled during the last 150 km of ascent. The counting rates on descent were significantly lower than the counting rates obtained at the same altitudes during ascent.

Figure 3 is a plot of radiation flux in counts/cm²/sec as a function of altitude in kilometers. Radiation flux was determined by dividing counts per minute $\times 1.6$ by the geometric factor (4.34 cm^2) \times efficiency of Geiger tube (0.85) \times time in seconds (60). Since only energetic particles would be counted from a 1.5π steradian look-angle with a mass greater than 20 g/cm^3 the counts per minute were multiplied by 1.6 to obtain the approximate total radiation flux.

Figure 4 is a plot of radiation flux in counts/cm²/sec as a function of altitude in kilometers.

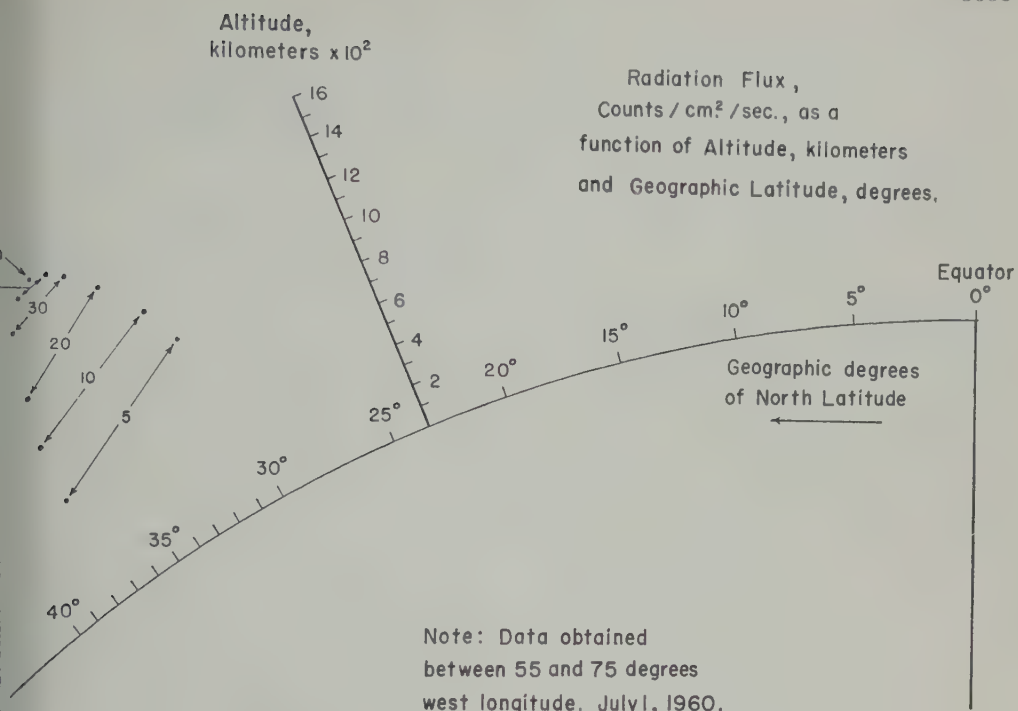


Fig. 4. Radiation flux as a function of altitude and geographic latitude.

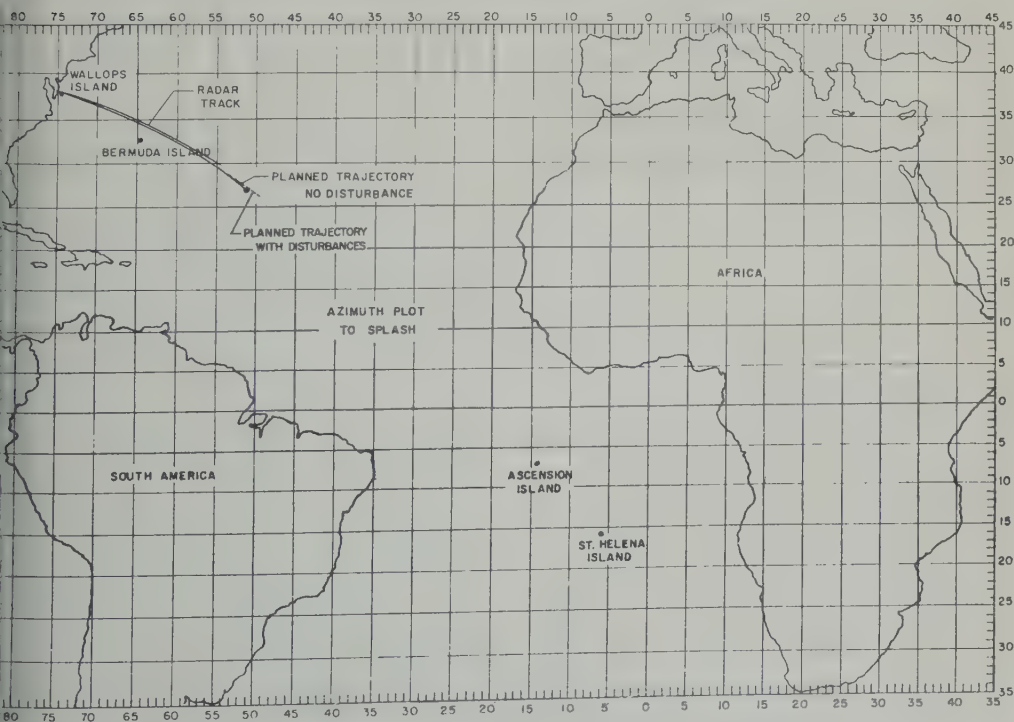


Fig. 5. Flight path of vehicle.

per square centimeter per second as a function of altitude and geographic latitude. The lines connect the two positions in space where a constant counting rate was obtained. The flux values measured near maximum altitude fall in between the flux measurement obtained from detectors *C* and *D* used by Van Allen on Explorer IV [Van Allen, McIlwain, and Ludwig, 1959].

Figure 5 shows the flight path of the vehicle. Maximum altitude was obtained at a geographic latitude of 34°N and a geographic longitude of

62° W. The calculated range of the vehicle was 2575 km.

In analyzing the data from the flight, there was no indication of malfunctioning of the radiation sensor during the 21-minute flight.

REFERENCES

- Van Allen, J. A., C. E. McIlwain, and G. H. Ludwig, Radiation observations with satellite Explorer IV, *J. Geophys. Research*, 64, 271-286 (1959).

(Received May 16, 1961; revised June 23, 1961)

Rocket Measurement of a Daytime Electron-Density Profile up to 620 Kilometers

J. E. JACKSON AND S. J. BAUER

*Space Sciences Division, Goddard Space Flight Center, NASA
Greenbelt, Maryland*

On April 27, 1961, a four-stage research rocket was fired from Wallops Island, Virginia, to measure the ionospheric electron-density distribution by means of Seddon's [1953] CW propagation technique. This experimental technique is based upon the measurement of the dispersive Doppler effect at two harmonically related frequencies—in this instance, $f = 12.267$ Mc/s and $6f = 73.6$ Mc/s. The actually measured quantities in the CW propagation experiment are the beat frequencies due to the difference between the received high- and low-frequency signals, the latter multiplied by the factor 6 at ground, for the two magnetoionic components.

These beat frequencies can be expressed by

$$f_{o,x} = \frac{6f}{c} \left[(n^{(h)}_{o,x} - n^{(l)}_{o,x}) \dot{r} + \int_0^R \frac{d}{dt} (n^{(h)}_{o,x} - n^{(l)}_{o,x}) dr \right] \pm \text{roll correction} \quad (1)$$

where c is the velocity of light in vacuo, $n^{(h)}$ the refractive index at the high frequency (73.6 Mc/s), $n^{(l)}$ that at the low frequency (12.267 Mc/s), and the subscripts o and x refer to the ordinary and extraordinary modes, respectively, and \dot{r} is the rocket velocity in the direction.

The first difference term in equation 1 represents a measure of the local electron density at the rocket, whereas the second term constitutes a correction term due to the variation of the electron-density distribution along the ray path. In a quiet ionosphere, the second term depends only on the geometry of the trajectory.

The present analysis was based on the sum of the ordinary and extraordinary beat frequencies which may be expressed by

$$f = \frac{6f}{c} \dot{r} [(n_o^{(h)} + n_x^{(h)}) - (n_o^{(l)} + n_x^{(l)})] \quad (2)$$

The sum F_o of the ordinary and extraordinary beat frequencies has the advantage of being free of roll effects, since the individual ordinary and extraordinary beat frequencies have equal but opposite roll corrections. Furthermore the quantity F_o is almost independent of the variation of the earth's magnetic field over the altitude and electron-density range covered by this experiment.

From F_o an apparent electron density corre-

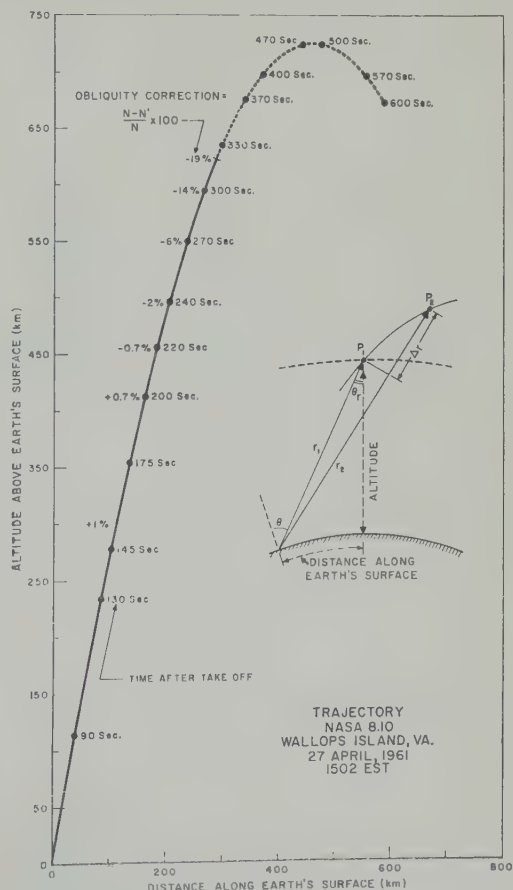


Fig. 1. Rocket trajectory.

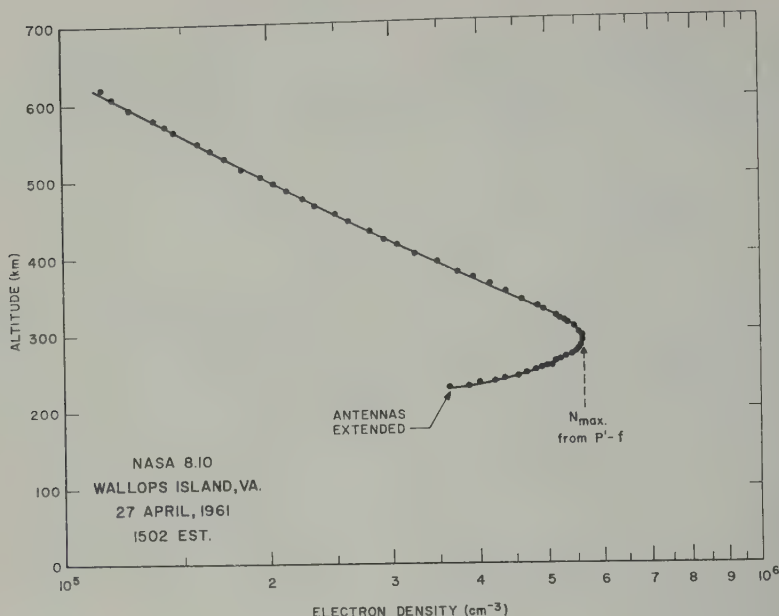


Fig. 2. Electron-density profile measured by rocket-to-ground CW propagation technique.

sponding to

$$N' = \frac{1}{\Delta r} \left(\int_0^{P_2} N dr - \int_0^{P_1} N dr \right) \quad (3)$$

can be obtained by means of the Appleton-Hartree formula together with trajectory information giving the radial velocity component.

The trajectory for this particular rocket flight is given in Figure 1. As the result of the curvature of the trajectory, the electron-density distribution along the ray path changes as the rocket moves from P_1 to P_2 (see insert in Fig. 1).

Ground-based ionosphere-soundings obtained simultaneously at Wallops Island, Virginia, Patuxent River, Maryland, and Fort Belvoir, Virginia, showed that the ionosphere was very quiet during the rocket flight and that there were no significant horizontal gradients in the electron density.

Thus, the actual electron densities N were derived from the apparent densities N' by using the following correction procedure

$$N = N' + \frac{\epsilon}{1 - \epsilon} (N' - \bar{N}) \quad (4)$$

with $\epsilon = r\theta/\dot{r} \tan \theta_r$, where θ is the zenith angle of the position vector at the receiving site and θ_r is the angle at the rocket, r is the radial

distance of the rocket with respect to receiving site and

$$\bar{N} = \frac{1}{r} \int_0^r N dr \cong \frac{1}{r} \int_0^r N' dr$$

In calculating \bar{N} , the integration below 620 km was based on electron-density data from ground-based vertical-incidence soundings. The magnitude of the obliquity correction is indicated on the trajectory curve in Figure 1, and it is seen that the maximum correction occurring at 620 km is -19 per cent, whereas below 400 km the correction is of the order of 1 per cent.

Figure 2 shows the resulting electron density profile up to an altitude of 620 km. The measurements began at an altitude of 225 km after ejection of the rocket nose-cone and extension of the rocket-borne transmitting antennas.

The experimental points represent averages over a time period of 1 second and were read at approximately 1-second time intervals up to an altitude of about 320 km, and at about 5-second intervals for the remainder of the profile.

The accuracy in reading the beat-frequency records from which electron densities were calculated is of the order of 1 per cent; the estimated accuracy of the trajectory determination is of the same order.

the maximum electron density as derived from the vertical-incidence sounder at Wallops Island, Virginia, is shown in Figure 2 to be in excellent agreement with the value derived from the rocket measurement. The electron-density distribution above the F_2 peak exhibits a practically constant logarithmic slope that is representative of a diffusive-equilibrium distribution in an isothermal ionosphere. This slope yields a height for the electron-ion gas of

$$H' = \frac{k(T + T_e)}{m_i g} = 200 \pm 11 \text{ km}$$

where k is the Boltzmann constant, T is the ionospheric (neutral particle) temperature, T_e is the electron temperature, m_i is the mean ionic mass, and g is the acceleration of gravity. Assuming thermodynamic equilibrium and a mean atomic mass ($m_i = 16$) corresponding to oxygen, we infer for the altitude region between 400 and 600 km a constant gas temperature of $1640 \pm 90^\circ\text{K}$.

This temperature inferred from our daytime profile is in good agreement with daytime temperatures derived from satellite density data, which also indicate appreciable diurnal and seasonal temperature variations [Kallmann-Bijl, 1961; Martin, Neveling, Priester, and Roemer, 1961]. Such a temperature variation is also apparent when our daytime (1502 EST) temperature of $T = 1640^\circ\text{K}$ is compared with an evening (2044 EST) temperature of $T = 1240^\circ\text{K}$ inferred from a positive ion-density profile above the F_2 peak measured at Wallops Island, Virginia, in November 1960 [Hanson and McKibbin, 1961].

Our temperature is also in agreement with the daytime value of electron temperature ($1800 \pm 300^\circ\text{K}$ at 1700 EST) directly measured on the Explorer VIII satellite [Bourdeau, Donley, Serbu, and Whipple, 1961]. The fact that our daytime temperature agrees with the neutral gas temperature derived from satellite density data as well as with the directly measured electron temperature is a good indication of thermodynamic equilibrium in the upper ionosphere.

Acknowledgments. The instrumentation used for this experiment has been designed by G. H. Spaid and J. R. Hagemeyer. The contributions of J. C. Seddon to the design of the system are gratefully acknowledged.

REFERENCES

- Bourdeau, R. E., J. L. Donley, G. P. Serbu, and E. C. Whipple, Jr., Measurements of sheath currents and equilibrium potential on the Explorer VIII satellite, AAS Symposium, March 17, 1961, Washington, D. C., *J. Am. Astronaut. Soc.*, 1961.
- Hanson, W. B., and D. D. McKibbin, An ion-trap measurement of ion concentration profile above the F_2 peak, *J. Geophys. Research*, **66**, 1667-1671, 1961.
- Kallmann-Bijl, H. K., Daytime and nighttime atmospheric properties derived from rocket and satellite observations, *J. Geophys. Research*, **66**, 787-796, 1961.
- Martin, H. A., W. Neveling, W. Priester, and M. Roemer, Model of the upper atmosphere from 130 km through 1600 km derived from satellite observation, COSPAR, Second International Space Science Symposium, Florence, Italy, April 13, 1961.
- Seddon, J. C., Propagation measurements in the ionosphere with the aid of rockets, *J. Geophys. Research*, **58**, 323, 1953.

(Received June 12, 1961.)

The Total Electron Content through the Entire Ionosphere by the Faraday Rotation Technique

O. BURKARD

Ionospheric Station, University of Graz, Austria

The polarization twist imposed on VHF radio waves by the ionosphere has been investigated by many workers using the moon as a reflector, the purpose being to determine the total electron column density. The difficulty in using only one frequency is the impossibility of an exact determination of the total electron content because the Faraday rotation is ambiguous in multiples of π . But I have shown recently that at daytime the electron content n_0 should vary with $\sqrt{\cos \chi}$, when the ionosphere is undisturbed, where χ means the zenith angle of the ionizing solar radiation. Plotting the Faraday rotation against $\sqrt{\cos \chi}$ is not a result of trying to find or proving a good agreement, but it is the result and a strong consequence of my new

model for the F layer [Burkard, 1961]. (It is surprising that the critical frequencies f_oF_2 do not show a similar correlation with $\sqrt{\cos \chi}$, because the latter values depend on the temperature in the F layer, and vary with time of day.) Hill and Dyce [1960] could not give values for n_0 using only one frequency, but have plotted the day-time values found by them against $\sqrt{\cos \chi}$ (Fig. 1). The result is as expected—a straight line. Clearly, absolute values for the electron content n_0 are now obtained by shifting this straight line vertically such a way that for $\cos \chi = 0$ also n_0 becomes zero. In such cases as in the early morning, however, when $\cos \chi$ is very small, we must use this $\cos \chi$ because we must not neglect the curvature of the earth. But there is no difficulty in finding the 'effective values' of $\cos \chi$ for 5 a.m. hours in the morning, if sufficient observational data of later morning hours are at hand.

Now a re-examination of the values N_mF_2 given by Dyce [1960] is also possible. Using observations from September 16, 1957, Figure 2 by Hill and Dyce is to be replaced by the

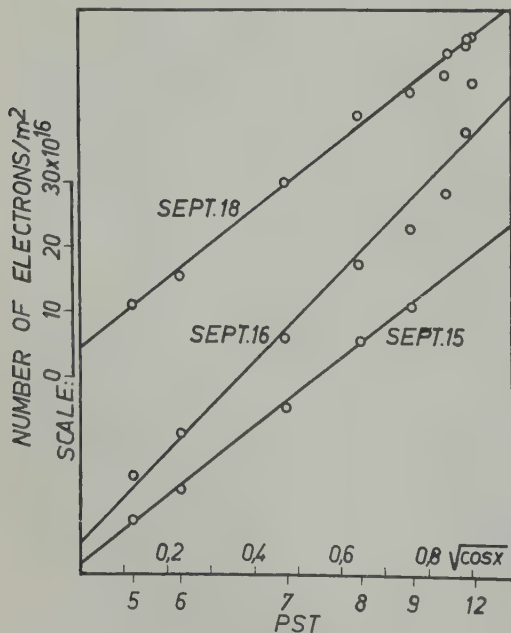


Fig. 1. Plot showing change of electron content with $\sqrt{\cos \chi}$ for several days in September 1957, observed by Hill and Dyce.

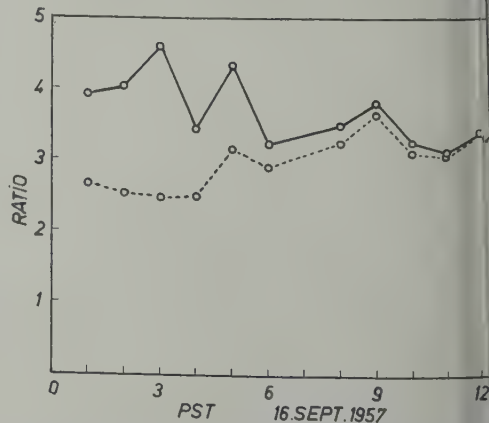


Fig. 2. Ratio of total electron content to electron content below the F peak. Dotted line, Hill and Dyce; solid line, recomputed values.

ing Figure 2. From it one can see that these computed values at night are somewhat lower than those given by Dyce. Therefore, a difference exists between the values observed by Hill and Dyce and those observed by Evans [1959], respectively by Bauer and Daniels [1959]. In all cases one finds the ratio N_a/N_b at night higher than at noon.

Finally, I wish to draw attention to the fact that the above-mentioned relation gives us a possibility of determining the variations of the ionizing solar radiation. The slope of the straight line in Figure 1 varies only with the varying intensity of solar radiation and is, therefore, independent of temperature in the ionosphere and other properties of it.

REFERENCES

- Bauer, S. J., and F. B. Daniels, Measurements of ionospheric electron content by the lunar radio technique, *J. Geophys. Research*, **64**, 1371-1376, 1959.
- Burkard, O., Studies of the new ionospheric model II, *Geofisica pura e appl.* **48**, 57-84, 1961.
- Dyce, R. B., Faraday rotation observations of the electron content of the exosphere, *J. Geophys. Research*, **65**, 2617-2618, 1960.
- Evans, J. V., The electron content of the ionosphere, *J. Atmospheric and Terrest. Phys.*, **11**, 259-271, 1957.
- Hill, R. A., and R. B. Dyce, Some observations of ionospheric Faraday rotation on 106.1 Mc/s, *J. Geophys. Research*, **65**, 173-176, 1960.

(Manuscript received May 20, 1961; revised June 24, 1961.)

Some Relationships between Short-Wave Fadeouts, Magnetic Crochets, and Solar Flares

LOREN W. ACTON

*Department of Astro-Geophysics, University of Colorado, and
National Bureau of Standards, Boulder, Colorado*

The purpose of this letter is to report the results of a study of correlations between the three classes of short-wave fadeouts (SWF) and the disk distribution and importance of the associated solar flares. We will also consider relationships between the type of SWF and the occurrence of simultaneous magnetic crochets (sometimes called 'solar flare effects' or SFE's).

Short-wave fadeouts are reported by the Central Radio Propagation Laboratory of the National Bureau of Standards, Boulder Laboratories, in its monthly tabulations of solar-geophysical data (F series, part B). They are arbitrarily divided into three classes according to the rapidity and general character of the recorded field strength decrease. The three types of fadeouts are illustrated in Figure 1 and described below.

S-SWF: sudden field strength decrease, taking from 1 to 5 minutes, depending on recorder time constant; gradual recovery.

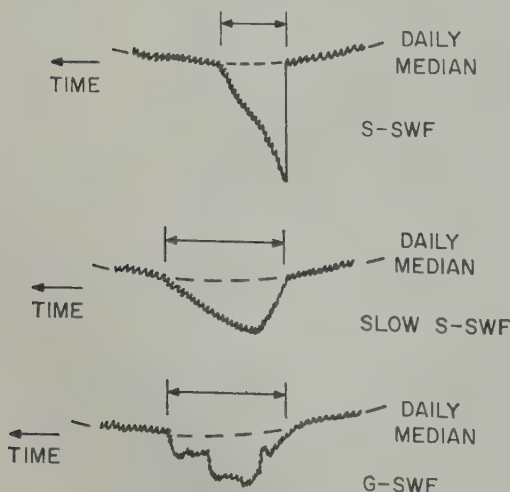


Fig. 1. Examples of field strength records of the three different types of short-wave fadeout.

Slow S-SWF: field strength decrease from 5 to 15 minutes but progresses slowly; gradual recovery.

G-SWF: gradual disturbance; fade irregular in both decrease and recovery.

In their study of limb flares and associated events *Warwick and Wood* [1959] suggest that there seemed to be a correlation between flare height and intensity and the type of associated SWF. In the present study several parameters have been investigated in an attempt to establish other relations associated with SWF type.

Figure 2 shows the distribution with latitude, measured from central meridian, of flares associated with each type of SWF. These SWF-associated flares occurred in 1958 and are listed in the CRPL *Solar-Geophysical* series. No east-west asymmetries were found, so in the final analysis the east and west values were combined for determining the variation with longitude.

Also included for comparison is the disk distribution of the IGY flares listed in the *Math-Hulbert Observatory Working Lists IGY Flares* [Dodson and Hedeman, 1960]. From the close similarity between the histogram of SWF-associated flares and that of all IGY flares, this indicates that there are no preferred longitudes for SWF-producing flares as a whole.

The disk distribution of the flares associated with each type of SWF was compared with the distribution for all IGY flares. There were no significant departures that were significant at the 5 percent level. However, the disk distribution of flares producing S-SWF's differs from that of all flares in the same way that the disk distribution of flares producing SFE's differs from that of all flares [Dodson and Hedeman, 1958].

The occurrences of associated magnetic crochets and SWF's have been examined for

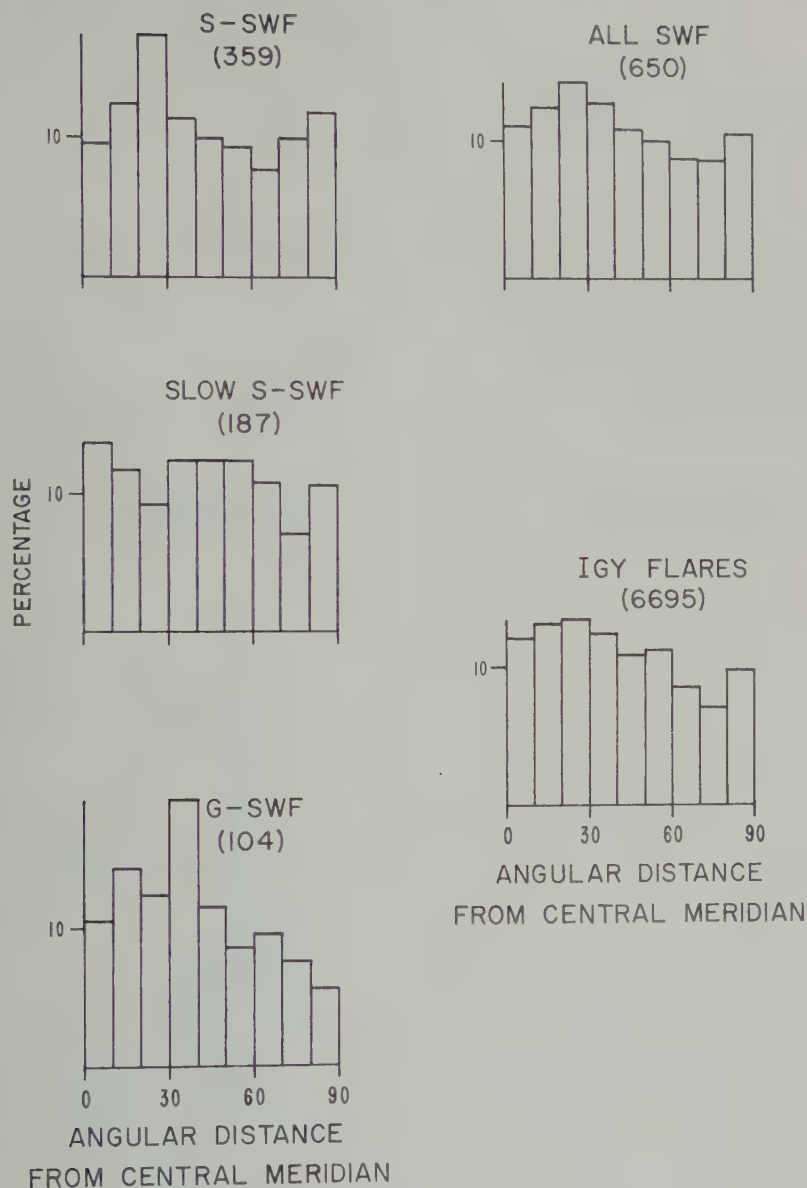


Fig. 2. Disk distribution of 650 SWF-associated flares that occurred in 1958, and disk distribution of all IGY flares listed in the *McMath-Hulbert Observatory Working List of IGY Flares*.

period January 1956 through December 1958. The data were taken from the 'National Data on Magnetic Disturbances' which is published regularly in this Journal. The results of the study are summarized in Table 3. It is evident from these curves that there is a strong preference for the occurrence of S-SWF's with SFE's. Note that the correlation holds at about the same level for each of

the 4 years of data examined. For the period studied it was found that

22.2 per cent of all S-SWF's were associated with SFE's.

7.7 per cent of all Slow S-SWF's were associated with SFE's.

3.1 per cent of all G-SWF's were associated with SFE's.

In view of the strong association shown between

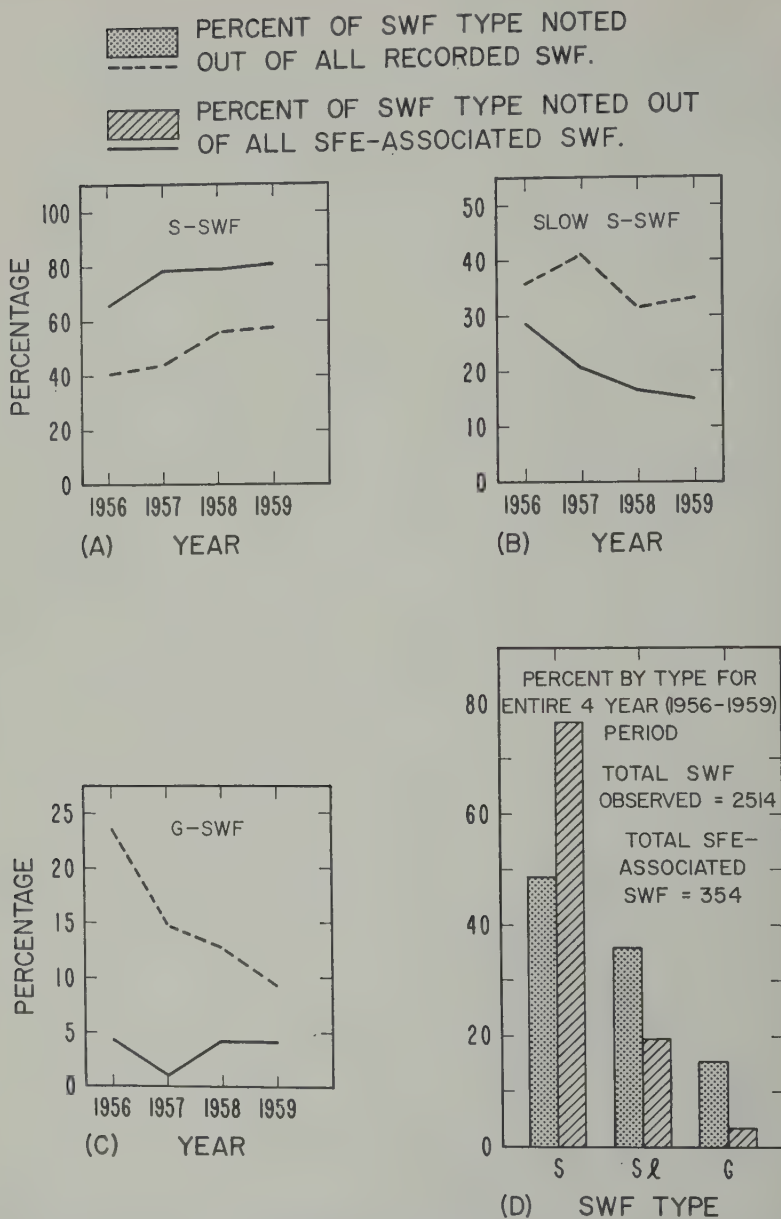


Fig. 3. Per cent by type of SFE-associated SWF compared with all SWF recorded from January 1956 through December 1959.

S-SWF's and SFE's it seems that the differences between the disk distribution of flares associated with S-SWF's and that of all IGY flares may be physically meaningful even though they are of small statistical significance.

We have examined, for the 650 SWF's discussed above, the dependence of SWF type on flare importance. Table 1 shows that there are

no appreciable differences in SWF type between fadeouts associated with flares of importance 2 and 3 and fadeouts associated with flares of importance less than 2. It was also found in this study that small flares (importance 1) which are associated with SWF's are six times more likely to produce an SFE than an 'average' small flare.

The H- α -flare patrol films of the Sacramento

TABLE 1. SWF Type as a Function of the Importance of the Associated Flare

are rtance	Per cent SWF by Type			Number of Fadeouts
	S-SWF	Slow S-SWF	G-SWF	
2	56.3	27.3	16.4	183
2	54.8	29.3	15.9	467

Observatory, for $3\frac{1}{2}$ months of the IGY, examined in an effort to establish a qualitative relation between optical flare appearance type of associated SWF. The study suggested that a compact, uniformly bright flare more likely to be associated with an S-SWF, that the more filamentary, nonuniform

flares could equally well produce any of the three SWF types.

Acknowledgment. While in the Department of Astro-Geophysics, University of Colorado, I was supported by a Ford Foundation fellowship.

REFERENCES

- Dodson, Helen W., and E. Ruth Hedeman, Crochet-associated flares, *Astrophys. J.*, **128**, 636-645, 1958.
- Dodson, Helen W., and E. Ruth Hedeman, *McMath-Hulbert Observatory Working List of IGY Flares*, IGY Solar Activity Report Series No. 12, IGY World Data Center A, Solar Activity, 1960.
- Warwick, C., and M. Wood, A study of limb flares and associated events, *Astrophys. J.*, **129**, 801-811, 1959.

(Received June 8, 1961.)

Airglow and the South Atlantic Geomagnetic Anomaly

K. D. COLE

*Antarctic Division, Department of External Affairs
Melbourne, Australia*

Dessler [1959] suggested why the Capetown anomaly of the geomagnetic field should explain the 'slot' between the inner and outer zones of trapped radiation. For present considerations this anomaly is regarded as a part of the large negative anomaly (here called the South Atlantic anomaly) which extends from the South China Sea southeast across Brazil and into the Indian Ocean off South Africa. One would expect a region of luminosity in the earth's upper atmosphere to be related to this feature of the geomagnetic field.

Now the observations of the spectral line 5577 Å in the airglow reported by Nakamura [1958] show high intensities in the range 15°–40° (geographic latitude). It is suggested here that these large values in the 5577 Å emission are partly due to the South Atlantic anomaly. It seems more than coincidence that the peaks of Nakamura's observations were obtained as the recording ship *Soya* passed through the South China Sea end and the South African end of this anomaly. Recent work [Huruhata and Nakamura, 1960] suggests a seasonal effect as well. This may be the result of seasonal variation in atmospheric density at mirror point heights.

A lower limit to the required flux of electrons (supposed here to be of average energy 50 kev) can be made. On the gross assumption that all the energy of an incident electron appears eventually as 5577 Å radiation, a lower limit flux of only $10^4 \text{ cm}^{-2} \text{ sec}^{-1}$ is required to explain

the mean airglow intensity of 240 rays found by Nakamura. Available fluxes of electrons in the 'slot' [Van Allen, 1959] are to be several orders of magnitude greater than this.

Traverses of the kind reported by Nakamura (perhaps performed from an aeroplane) repeated in other longitudes may clarify the relationship of this component of airglow to the structure of this geomagnetic anomaly. One would expect a pronounced longitude variation of intensity of this airglow line in the vicinity of 35°S, with a maximum over southeastern Brazil, the center of this anomaly. Such observations should complement the experiments suggested by Cladis and Dessler [1961] to determine the relationship of X rays from Van Allen electrons and the Capetown anomaly.

Cladis, J. B., and A. J. Dessler, X rays from the Van Allen Belt electrons, *J. Geophys. Research*, **66**, 343–350, 1961.

Dessler, A. J., Effect of magnetic anomaly on trapped particle radiation in the geomagnetic field, *J. Geophys. Research*, **64**, 713–715, 1959.

Nakamura, J., Latitude effect of night airglow, *Rept. Ionosphere Research Japan*, **12**, 419–420, 1958.

Huruhata, M., and J. Nakamura, Latitude dependence of airglow emission, *C.I.G. News*, **10**, p. 92, 1960.

Van Allen, J. A., The geomagnetically trapped particle radiation, *J. Geophys. Research*, **64**, 1683–1687, 1959.

(Received June 19, 1961.)

Possible Effects of Magnetic Field Alignment on Meteoric Ionization

WALTER F. BAIN

Page Communications Engineers, Inc.
Washington 7, D. C.

Field alignment of ionization along the earth's magnetic lines has been shown to result in enhanced scattering efficiencies, as in the aurora or observed in the equatorial *E* region by Bowles, Cohen, Ochs, Balsley [1960], among others. It is of interest to consider also the effect of such field alignment on ionization that is initially formed as a thin dense column, such as a meteor trail, parallel to the lines of magnetic force. The tendency of the magnetic field to maintain this initial alignment could to some extent offset normal diffusion processes and result in unusually long trail durations. Such durations have often been observed on overdense meteors, but never adequately explained.

The effect on underdense meteors would also tend to increase duration. The added requirement of specular reflection imposed on underdense trails, would, however, make the increased durations apparent only on radars beamed perpendicular to the magnetic lines of force. This enhancement in trail duration could, then, give the appearance of sustained radiant activity in a direction parallel to the lines of magnetic force on equipments sensitive to echo duration. In addition, it is possible that the peak echo amplitude could also be increased. This could result from a similar magnetic control of the mean-free-path, which has been shown by Manning [1958] to be a determining factor in initial trail radius and hence in initial density. These effects could modify radiant survey results in a manner dependent on the magnetic field angle at the measuring site.

A case in point which might be considered is

the considerable daylight radiant activity first delineated during the summer months in the later 1940's [Lovell, 1954]. Radiants in this region lie very close to the plane of the ecliptic which, during June and July, is at its maximum elevation over the southern horizon. At the latitude and magnetic-dip angle of Manchester, England, where these radiants were first observed, this configuration places them within a few degrees of being parallel to the lines of force of the earth's magnetic field. Thus, any general concentration of sporadic meteors in the plane of the ecliptic would manifest itself as an increase in activity during these months. The high degree of activity during this period, therefore, may not be entirely due to radiant concentration alone; it may be showing some degree of enhancement due to field-aligning effects on the sporadic trails originating in the ecliptic plane. It would, in fact, be of interest to consider whether the measurement of the ecliptic concentration itself might not have been influenced by this effect.

REFERENCES

- Bowles, K. L., R. Cohen, G. R. Ochs, and B. B. Balsley, Radio echoes from field-aligned ionization above the magnetic equator and their resemblance to auroral echoes, *J. Geophys. Research*, **65**, 1853-1855, 1960.
Lovell, A. C. B., *Meteor Astronomy*, chapter 28, XVIII, Oxford University Press, London, 1954.
Manning, L. A., The initial radius of meteoric ionization trails, *J. Geophys. Research*, **63**, 181-196, 1958.

(Received May 26, 1961.)

A Comparison of Theory and Observation of the Echo I Satellite

R. BRYANT

*Theoretical Division, Goddard Space Flight Center
National Aeronautics and Space Administration
Silver Spring, Maryland*

We have continued the study of the orbit of the Echo I satellite reported in an earlier communication [Jastrow and Bryant, 1960; Shapiro and Jones, 1960; Muhlemen, Hudson, Holdridge, Carpenter, and Oslund, 1960].

As reported earlier, the primary effect of solar radiation pressure on the Echo satellite is to change the eccentricity, and hence the altitudes of apogee and perigee. The period will also be changed when a portion of the orbit is in the shadow of the earth. The slow changes in period produced by the shadowing must be subtracted from the observed period changes, in order to reveal the net effect produced by drag.

Assuming neutral drag effects only, the value of atmospheric density near perigee altitude is inferred from the net rate of change of the orbital period after the changes due to shadowing have been subtracted. Since the perigee altitude descends from 1500 to 930 km and rises again to nearly 1500 km during the first year in the life of Echo I, we are able to sample the densities in this range of altitudes twice during that period. Thus we can attempt a separation of the time dependence from other effects in the atmosphere. The ability to sample the same altitudes at different times is a unique feature of Echo I among the satellites launched thus far.

Figure 1 is a plot of the perigee altitude of Echo I as determined from Minitrack interferometer observations, Smithsonian Astrophysical Observatory Baker-Nunn optical observations, and optical observations obtained from modified calibration cameras at the Minitrack stations. Each orbit correction (i.e., determination of the mean elements from the observations), uses observational data from a 4-day interval. The orbit correction program is based on the modified Hansen theory [Musen, 1959], and includes the effects of the second and fourth harmonics of the earth's gravitational potential.

The solid curve in Figure 1 is the theoretical perigee height obtained from a theory of radiation pressure (including shadowing effect) developed by Bryant [1961], combined with the perturbations produced by second, third, and fourth harmonics of the earth's gravitational potential, lunar and solar terms, and atmospheric drag. The theory does not allow for radiation re-emitted by the earth. Also, it assumes a simple exponential model of atmospheric density. Several other authors have developed comparable theories. Wyatt [1961] has considered the long-period effects of radiation pressure, Zadunaitsky, Shapiro, and Jones [1961] have given a complete analysis of their data on Echo I in which they have included the above perturbations plus the effect of the fifth gravitational harmonic and the re-emission of radiation from the earth.

The starting point for the theoretical curve in Figure 1 is taken as 26 days after launch, when it is believed that outgassing had been completed and the mass had stabilized. The residual in the perigee altitude is 7.5 km.

Figure 2 shows the rate of change of perigee of Echo I vs. time, as determined from the orbit computing program in use at the Goddard Space Flight Center. Also included in the figure is the theoretical curve giving the change in perigee produced by the shadowing effect. Energy may be removed from or added to the satellite by this effect, depending on the orientation of the orbit, and it is necessary to separate such energy changes from those produced by drag before conclusions can be drawn about atmospheric density. Figure 2 shows the total rate of change of period as observed, and the net rate of change after the shadowing effect has been removed. The gap in the data in Figure 2 during December is for a period between the cessation of the Minitrack transmitter and the receipt of routine optical observations from the Minitrack.

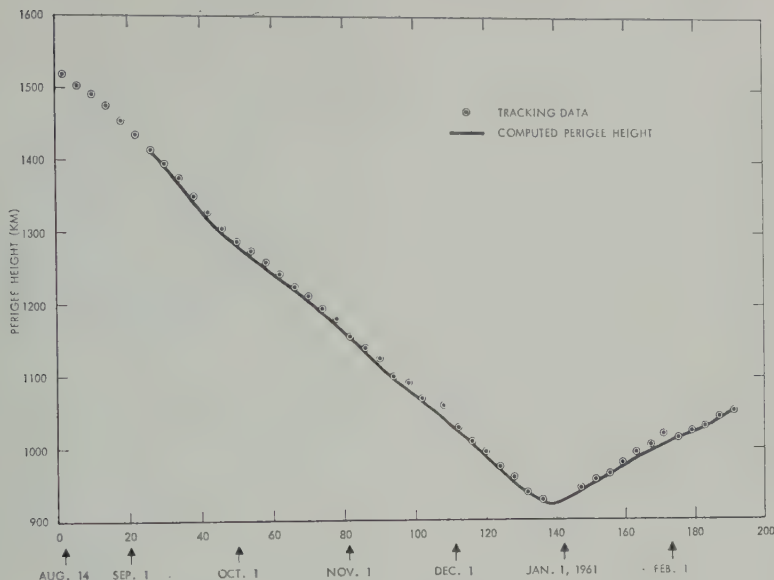


Fig. 1. Effect of radiation pressure on Echo I.

tions. The data given show pronounced peaks in the rate of change of period centered on the dates of November 14 and December 5. Major solar flares were observed at approximately these same times.

If one assumes neutral drag only, densities can be inferred from the rates of change of period. Figure 3 shows the densities obtained from the period changes of Echo I by application of the

King-Hele [1960] formula, which relates atmospheric density at half a scale height above perigee height to the rate of change of period. By iterative application of this formula we obtain an apparent scale height of 260 km.

It must be pointed out, in considering these density values, that the atmosphere cannot be considered as static at high altitudes. The diurnal effect can produce changes in the density of as

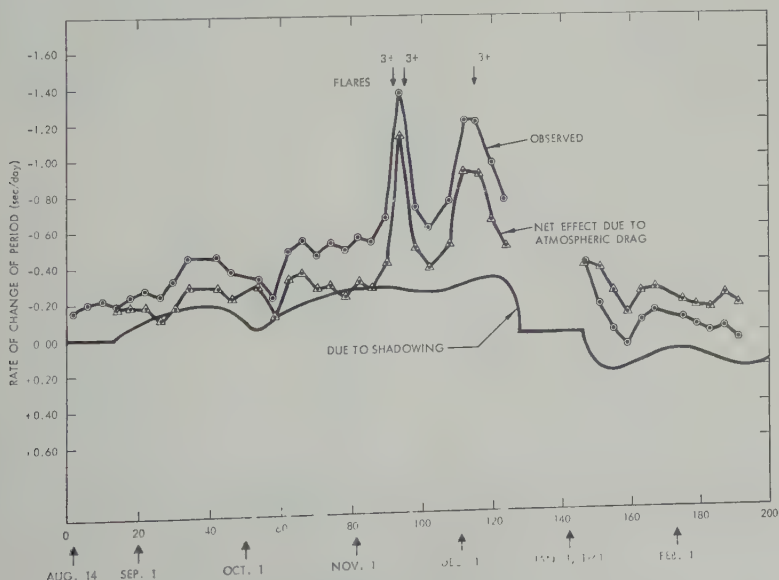


Fig. 2. Rate of change of period of Echo satellite.

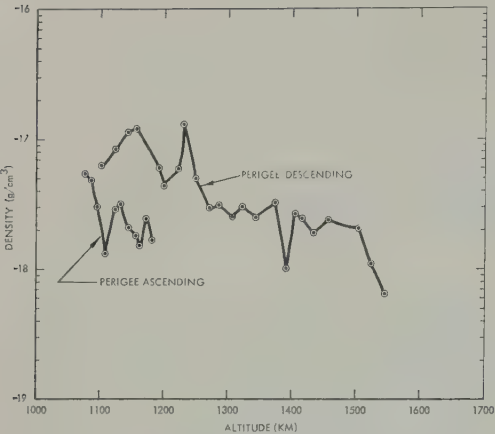


Fig. 3. Atmospheric densities obtained from Echo I.

much as a factor of 30, and the effects associated with solar activity can produce changes of as much as a factor of 4, and possibly much greater. The seasonal effect may be as much as a factor of 1.5 [Priester and Martin, 1960; King-Hele, 1960; Paetzold and Zschorner, 1960].

The diurnal effect could negate the validity of the formula for obtaining densities from rates of change of period, since the derivation of the formula is based on the assumption that most of the drag occurs in the vicinity of perigee. It is conceivable that a portion of the orbit other than the portion near perigee might be in the diurnal bulge, and that the drag force could be greater along that portion than along the portion near perigee. For the period covered in this study the local time at perigee changed from 18 hours to 11 hours, and it is considered, since the maximum diurnal effect is at about 14 hours, that the perigee portion of the orbit is near enough to the diurnal bulge to sustain the validity of the densities obtained by the formula.

Figure 4 shows the flux of solar radiation at 20 cm, which constitutes a good indication of the level of solar activity. According to Figure 4, solar activity showed a downward trend, from a flux of approximately 180 on August 14 to 100 on November 2. Jacchia [1960] has obtained an empirical relation between atmospheric density and the 20-cm flux by analysis of several satellite

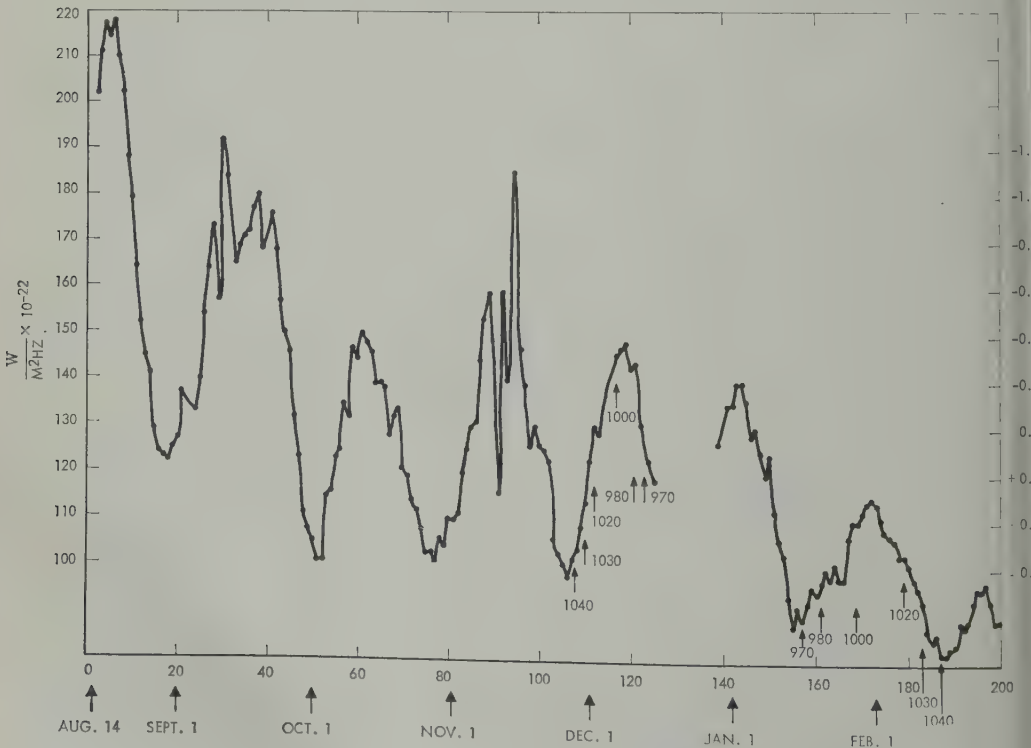


Fig. 4. Solar flux data.

its with perigee altitudes ranging between 1000 and 700 km. According to Jacchia, the density is proportional to the first power of the 20-cm flux. Priester has analyzed the Echo data and finds that the density is proportional to the square of the 20-cm flux at altitudes between 1000 and 1500 km. We find approximately the same relationship reported by Priester through our analysis of the density variations in Figure 3. The upper curve represents the density as perigee altitude is descending, and the lower curve that as perigee altitude is rising. Using altitudes of 1065, 1040, 1020, 985, 965 km, average flux values at the perigee times corresponding to each of these altitudes, and adjusting for the diurnal effect by using the Bonn density formula [Martin, Neveling, Priester, and Roemer, 1961], we obtain values of 1.5, 8.5, 1.8, and 2.0 for the power law exponent in the density-flux relation. The value 8.5 is obtained at the epoch of December 2 and is undoubtedly influenced by the flare activity.

Acknowledgments. I am indebted to Dr. Robert F. Fritter for a number of valuable discussions of this problem. I also wish to thank Robert Devaney and Arthur Smith, Jr., of the Theoretical Division of the Goddard Space Flight Center, who carried out the orbit corrections and programmed the necessary program for a large-scale computer.

REFERENCES

Fritter, R., The effect of solar radiation pressure on the motion of an artificial satellite, *NASA Tech. Note*, Goddard Space Flight Center, 1961.
 Jacchia, L. G., A variable atmospheric-density model from satellite accelerations, *J. Geophys. Research*, **65**, 2775-2782, 1960.
 Fritter, R., and R. Bryant, Variation in the orbit

- of the Echo satellite, *J. Geophys. Research*, **65**, 3512-3513, 1960.
 King-Hele, D. G., Method for determining the changes in satellite orbits due to air drag, *Space Research*, Proceedings of the First International Space Science Symposium, edited by H. K. Hallmann-Bijl, North Holland Publishing Company, Amsterdam, 1960.
 Martin, H. A., W. Neveling, W. Priester, and M. Roemer, Model of the upper atmosphere from 130 through 1600 km, derived from satellite orbits, *Mitteilung der Universitaets-Sternwarte, Bonn* nr. 35, 1961.
 Muhleman, D. G., R. H. Hudson, H. B. Holdridge, R. L. Carpenter, and K. C. Oslund, Observed solar perturbations of Echo I, *Science*, **132**, 1487, 1960.
 Musen, P., Application of Hansen's theory to the motion of an artificial satellite in the gravitational field of the earth, *J. Geophys. Research*, **64**, 2271-2279, 1959.
 Paetzold, H. K., and H. Zschorner, Bearings of Sputnik 3 and the variable acceleration of satellites, in *Space Research*, North Holland Publishing Company, Amsterdam, 1960, pp 24-36.
 Priester, W., and H. A. Martin, Solare und tageszeitliche Effekte in der Hochatmosphaere aus Beobachtungen kuenstlicher Erdsatelliten, *Mitteilung der Universitaets-Sternwarte Bonn*, nr. 29, Westdeutscher Verlag, Koln-Opladen, 1960.
 Shapiro, I., and H. M. Jones, Perturbations of the orbit of the Echo balloon, *Science*, **132**, 1484-1486, 1960.
 Wyatt, S. P., The effect of radiation pressure on the secular acceleration of satellites, *Smithsonian Astrophysical Observatory Special Rep.* No. 60, 1961.
 Zadunaisky, P. E., I. Shapiro, and H. M. Jones, Experimental and theoretical results on the orbit of Echo I, *Smithsonian Astrophysical Report*, No. 61, 1961.

(Received June 19, 1961; revised June 23, 1961.)

Recalculations of Cloud Electrification Based on a General Charge-Separation Mechanism

J. D. SARTOR

The Rand Corporation, Santa Monica, California

Professor B. J. Mason, Imperial College, London, in private correspondence has made several important comments and corrections concerning the paper 'Calculations of Cloud Electrification Based on a General Charge-Separation Mechanism' by Sartor [1961]. Since the outcome is favorable to my hypothesis, I hasten to bring the adjusted results to the attention of the reader. The items requiring correction are:

1. The factor q_0 of Table 1 is dimensionless and refers to values of the electric field, E , in volts per centimeter, not esu as indicated.

2. The E should not appear in the right-hand term of equation 6, and the numerical factor should be $2\pi^2$, not $4\pi^2$.

3. The values of q_0 in Table 1 should be multiplied by 300 when entering equation 6, or, more precisely, equation 6 should contain an extra factor of 300 in the exponent and q_0 entered as given in Table 1.

Professor Mason points out that if the field is allowed to vary in computing the volume charge density, equation 2, the field growth, equation 6, takes on a different form in which the exponent contains t in place of t^2 , and the double summation is replaced by its square root. This suggested generalization and a clarification of the summation notation have been incorporated into a rederivation of the field growth equation along the same general lines given in the paper, but considering the field E to vary with time in the computation of the volume charge density.

The revised growth equation becomes

$$\frac{E}{E_0} = \exp \left\{ \left[1200\pi^2 \epsilon q_0 \sum_{j=2}^k \sum_{i=1}^{j-1} n_i r_i^2 (V_j - V_i)^2 n_i r_i^2 + \frac{\lambda^2}{4} \right]^{1/2} t + \frac{\lambda}{2} t \right\}$$

where the notation is as before. The summation notation applies to the drop-size distribution divided into k intervals in which each interval

is characterized by a mean radius r and concentration n , and the k intervals are run by increasing radius from $j = i = 1$ through $j = i = k$.

The computations of the field growth based on the revised equation are shown in Figure 1. The charge-exchange coefficient is taken as $\frac{1}{3} \times 10^{-2}$, a value slightly less than the minimum expected value. The previous value was 1 corresponding closely to the maximum expected value of 1.08×10^{-2} . The results are shown parametrically according to the value of ϵ , particle-interaction efficiency.

As before, it appears that, assuming reasonable or even low particle-interaction efficiencies, the proposed general charge-separation mechanism is capable of building fields of thunderstorm magnitude in the required time, provided

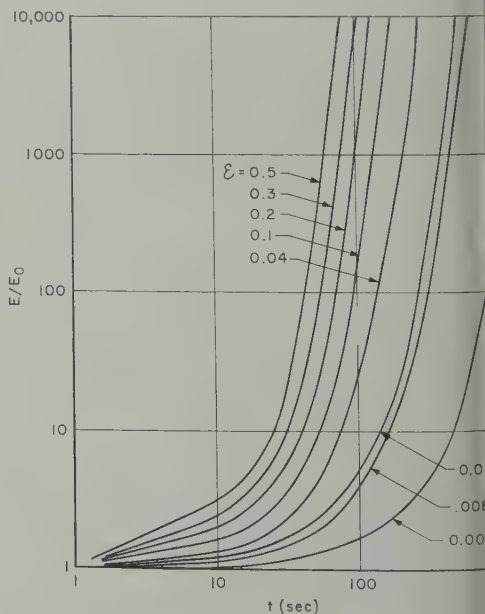


Fig. 1. Field enhancement vs. time for various values of ϵ shown beside curve.

appropriate cloud- and rain-drop distributions present.

Acknowledgments. The corrections pointed out by Professor B. J. Mason are profoundly appreciated, and his suggestions for improvement of the field enhancement equation gratefully acknowledged.

REFERENCES

- Sartor, J. D., Calculations of cloud electrification based on a general charge-separation mechanism, *J. Geophys. Research*, 66, 831-838, 1961.

(Received June 20, 1961.)

Comment on Paper by B. Le Méhauté, 'A Theoretical Study of Waves Breaking at an Angle with a Shore Line'

P. GROEN

*Division of Oceanography
Koninklijk Nederlands Meteorologisch Instituut
De Bilt, Netherlands*

With reference to the paper by *Le Méhauté* [1961], it may be worth while to draw the attention of your readers to the fact that this subject was dealt with many years ago by *Groen and Weenink* [1950].

finding breaker characteristics along a straight coast, *Trans. Am. Geophys. Union*, *31*, 398-399, 1950.

Le Méhauté, B., A theoretical study of waves breaking at an angle with a shore line, *J. Geophys. Res.*, *66*, 495-499, 1961.

REFERENCES

Groen, P., and M. P. H. Weenink, Two diagrams for

(Received April 24, 1961.)

Author's Reply to the Preceding Comment

B. LE MÉHAUTÉ

*Department of Civil Engineering
Queen's University at Kingston, Ontario, Canada*

I thank Dr. P. Groen very much for completing my bibliography by his interesting paper. There are indeed great similarities between the two studies and fairly good agreement between the practical results. Discrepancies come from the fact that breaking conditions are expressed slightly differently, since I have used the theory

of Miche valid for any depth instead of some theoretical graphs, partly based on the solitary wave theory.

I do not think that these theoretical differences are of great practical importance.

(Received June 9, 1961.)

A Note on Ocean Sediment Thickness from Surface Wave Dispersion¹

ROBERT L. KOVACH AND FRANK PRESS

*Seismological Laboratory
Pasadena, California*

Oceanic Rayleigh wave dispersion data have often been used as a means of estimating the thickness of unconsolidated sediments on the ocean floor. The difference between the thickness of the fluid layer determined from the dispersion data and the depth of the water is taken as the thickness of the sediments. Recently, *Evison, Ingham, Orr, and Le Fort* [1960] raised some questions about the validity of this method. They used oversimplified theoretical dispersion models for which crustal thickness and fluid depth were related by a ratio and stated that the value of sediment thickness obtained is wholly dependent on the assumed ratio of crustal thickness to fluid depth.

First, it must be pointed out that we are dealing with the steep portion of the group velocity curve where the observed period of Rayleigh waves can be determined very precisely.

Second, it should be emphasized that the observed data should be in the velocity range for which the water and sediments are the controlling factors, namely, 1.5 km/sec to about 3 km/sec. None of *Evison's* data were in this range. In Figure 1 are shown three dispersion curves of Rayleigh waves for an oceanic path calculated on the IBM 7090. The velocity parameters are based on *Raitt's* [1956] refraction data in the Pacific Ocean and *Gutenberg's* [1959] velocity structure for the mantle. Most of the refraction data at sea show a fairly uniform upper crustal layer, 1 to 2 km thick, and an intermediate crustal layer varying from 3 to 9.5 km in thickness. For these calculations we have selected as the variable the thickness of the intermediate crustal layer. As can be seen in Figure 1, in the velocity range of 1.5 to about 3 km/sec, the effect of crustal thickening is indeed small. Further, if the dispersion method is restricted to ocean basins, as it should be since margins are associated with other sedi-

¹ Contribution 1038, Division of Geological Sciences, California Institute of Technology.

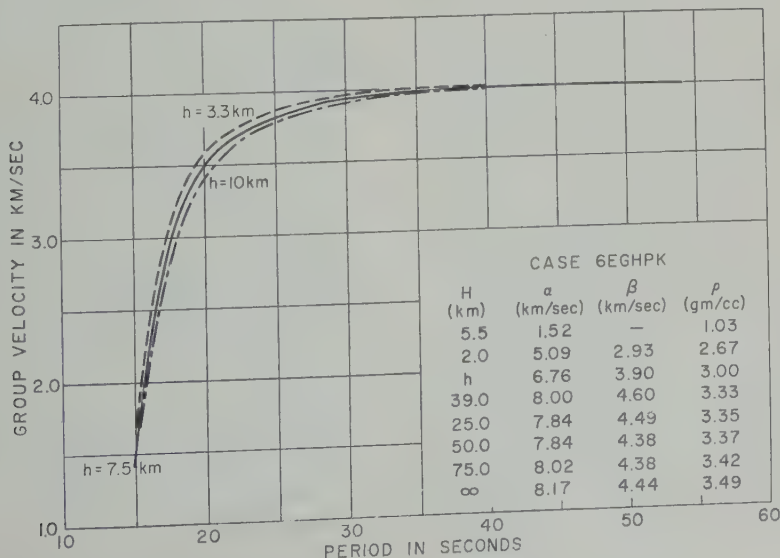


Fig. 1. Oceanic Rayleigh wave dispersion showing effect of crustal thickening.

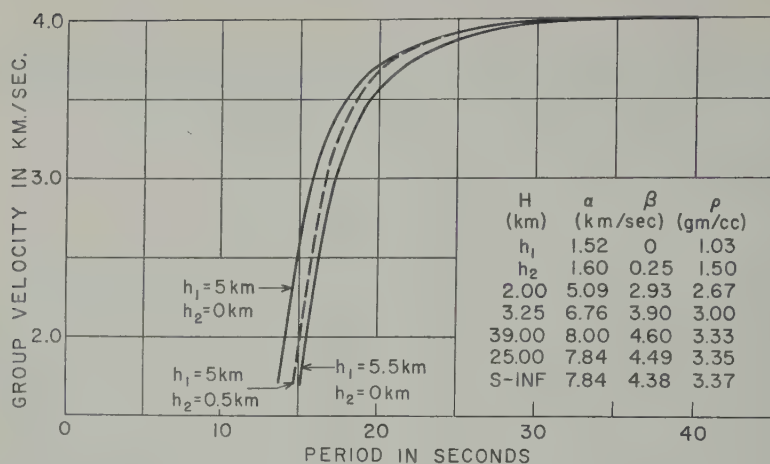


Fig. 2. Oceanic Rayleigh wave dispersion showing effect of unconsolidated sediment layer.

mentary processes and with anomalous crustal thickness, then a value close to 5 km is always found [Raitt, 1956; Ewing and Ewing, 1959] for the thickness of the crust. Thus, crustal thickness is not a significant factor in determining unconsolidated sediment thicknesses.

By far the most important factor is the small but finite rigidity of the sediments. Theoretical calculations for a sediment and a no-sediment case are presented in Figure 2. The two solid curves are for a 5-km and a 5.5-km fluid layer with the unconsolidated sediment assumed to be combined in the fluid layer and the dashed curve is for a 5-km water layer and a 0.5-km sediment layer of low rigidity. The velocity parameters selected for the sediment layer are based on the most recent data available [Fry and Raitt, 1961; Oliver and Dorman, 1961]. On the basis of the theoretical curves shown here, the error in sediment thickness would be 0.1 km if the solidity of the sediments were neglected. Allowing for the solidity of the sediments will certainly reduce the error in determinations of sediment thickness. The range of sedimentary thickness found by dispersion analyses is 0.5 to 1 km, whereas the typical range found by refraction measurements is 0.2 to 1 km. These determinations of sedimentary thicknesses have certainly been of the right order of magnitude. Other sources of error will be more important. For example, it is difficult to estimate the

uncertainty in the surface wave method because of irregular topography and contributions from nonbasin portions of the path. When dispersion data and refraction data are available for the same region we will be better able to estimate the ultimate precision of the method.

Surface wave methods are not meant to be competitive with explosion methods. The former are the only source of information in remote oceans not covered by refraction surveys.

REFERENCES

- Evison, F. F., C. E. Ingham, R. H. Orr, and J. E. Le Fort, Thickness of the earth's crust in Antarctica and the surrounding oceans, *Geophys. J. S.*, 289-306, 1960.
- Ewing, J., and M. Ewing, Seismic refraction measurements in the Atlantic Ocean basins, in the Mediterranean Sea, on the mid-Atlantic ridge, and in the Norwegian Sea, *Bull. Geol. Soc. Am.*, 70, 291-318, 1959.
- Fry, John C., and R. W. Raitt, Sound velocities at the surface of deep sea sediments, *J. Geophys. Research*, 66, 589-597, 1961.
- Gutenberg, B., The asthenosphere low-velocity layer, *Ann. geofis. (Rome)*, 12, 439-460, 1959.
- Oliver, J., and J. Dorman, Nature of oceanic seismic surface waves with predominant periods of 6 to 8 seconds, *Bull. Seism. Soc. Am.*, 1961.
- Raitt, R. W., Seismic refraction studies of the Pacific Ocean basin, *Bull. Geol. Soc. Am.*, 67, 1623-1640, 1956.

(Received May 28, 1961.)

On the Geochemical Character of Iodine in Meteorites

GORDON G. GOLES¹ AND EDWARD ANDERS*Enrico Fermi Institute for Nuclear Studies and Department of Chemistry
University of Chicago, Chicago, Illinois*

The recent detection of excess Xe^{129} in chondrites [Reynolds, 1960a, b; Zähringer and Ventner, 1961] is of great importance in establishing the chronology of the early solar system [Goles and Anders, 1960; 1961]. In order to interpret these results in terms of meaningful I^{129} - Xe^{129} decay intervals, however, one must investigate two problems in particular: Is the Xe^{129} concentrated in the same mineral phases as the iodine? What effect will diffusion losses of Xe^{129} have on the decay intervals calculated [Goles, Fish, and Anders, 1960]? A step toward answering both questions is the investigation of the geochemical character of iodine in meteorites. The program of iodine analyses by neutron activation, preliminary results of which were reported previously [Goles and Anders, 1960], has been continued in order to study the distribution of iodine among meteoritic phases. The analytical procedures used were essentially the same as those described in our previous paper [Goles and Anders, 1960].

Three types of experiments have been performed: leaching experiments; abundance determinations on troilite and metal phases from iron meteorites; and analysis of mechanically separated fractions of a chondrite.

In the leaching experiments, the irradiated sample, in the form of fragments weighing ~ 10 to ~ 500 mg, was treated with a hot ($\sim 90^\circ\text{C}$) ammonia-buffered solution of pH ~ 8 for 15 minutes. After this comparatively mild treatment, the supernate was decanted and filtered, the residue washed with distilled water, and the soluble and insoluble fractions were worked up separately. Table 1 presents the data on these experiments. There is a marked difference in the behavior of I^{127} and fission-product iodine in the two classes of meteorites represented. Little or no fission-product iodine is leachable from

the bronzite and hypersthene chondrites Richardton or Bruderheim, whereas significant fractions of these activities are found in the leaching solutions from the carbonaceous chondrite Murray. Although this may indicate that uranium in Murray resides partly in a water-soluble phase, it may also mean that the uranium-bearing grains of this meteorite are small enough so that some of the fission fragments recoil out and come to rest in water-soluble materials. In contrast, a large fraction of the I^{127} was found in the leaching solution in all four experiments. In each case, a somewhat smaller relative amount of I^{131} (from the decay of Te^{131}) accompanied the I^{127} . We interpret these observations as indicating that most of the iodine in Richardton and Bruderheim, and at least one-third that in Murray, resides in one or more water-soluble phases.

The experiments on separated phases from iron meteorites indicate that iodine is accommodated more easily in troilite than in metal. The metal phases of Canyon Diablo and Toluca meteorites contain ~ 30 and ~ 250 ppb iodine, whereas the troilite phases contain ~ 60 and ~ 1000 ppb iodine, respectively. Thus, although there is about a tenfold variation in the iodine abundances of these two meteorites, the iodine is decidedly chalcophilic in both cases. Similar variations of trace element abundances were found by Reed, Kigoshi, and Turkevich [1960], who have shown that Toluca troilite is richer in Bi, Tl, and probably Hg, than is Canyon Diablo troilite. This trend is reversed in the case of Te, however: Canyon Diablo troilite contains ~ 5 ppm tellurium, and Toluca troilite, ~ 1.7 ppm tellurium. The metal phases contain ~ 90 and ~ 50 ppb tellurium, respectively, implying that this element is more strongly chalcophilic than is iodine.

From this information, and from the evidence supplied by the leaching experiments, we hypothesize that the iodine and a portion of the

¹ Now at the Department of Chemistry, University of California, La Jolla, California.

TABLE 1. Results of Leaching Experiments

Sample	Total			% in Leach Fraction*		
	I, ppb	Te, ppm	U, ppb	I ¹²⁸	I ¹³¹	Fission
Richardton III	33 ± 4	0.52 ± 0.03	13 ± 2	~54	~27	~2
Bruderheim I	16 ± 6	0.46 ± 0.03	11 ± 1	~60	~52	<1
Murray III	300 ± 40	1.69 ± 0.10	27 ± 3	~35	~13	~24
Murray IV	230 ± 30	1.84 ± 0.11	15 ± 2	~39	~29	~26

* The relative activities of I¹²⁸, I¹³¹ and fission-product iodine found in the leach fraction probably represent roughly equivalent fractions of I, Te, and U initially in water-soluble phases.

tellurium in Richardton and Bruderheim reside in oldhamite (CaS) or a similar phase (e.g., MgS). Oldhamite is one of the chief water-soluble accessory minerals in chondrites, the only other common one being lawrencite (FeCl₂) [Buddhue, 1941]. The chalcophilic character of iodine in iron meteorites leads us to believe that the sulfide, oldhamite, is more likely to be the host mineral of iodine in chondrites than is lawrencite. This inference, though based entirely on circumstantial evidence, is also supported by observations of extensive crystal imperfections in meteoritic MgS reported by *DuFresne and Anders* [1961], and by the fact that terrestrial chloride minerals are normally depleted in iodine, relative to the I/Cl ratio in the earth's crust [Goldschmidt, 1954, pp. 585, 604]. Of course, this conclusion should be subjected to further experimental tests, especially since preliminary data on the iron meteorite Sardinia indicate that this meteorite, with a high lawrencite content, also happens to be high in iodine.

There is further circumstantial evidence to support the identification of oldhamite as the iodine host mineral. *Tassin* [1908] has reported enrichment of Ca and S in a fine fraction of the Allegan chondrite and a comparable enrichment of MgS has been found in a fine fraction from the Abee chondrite [DuFresne and Anders, 1961]. We have performed analyses on mechanically separated fractions of the Bruderheim chondrite, comparing the iodine content of a fine fraction (minus 325 mesh) with that of the whole and of a sample containing predominately large fragments. Although the precision of these analyses was not very good, due to interference by fission-product iodine activities, there seems to be some enrichment of iodine in the fine fraction. If Bruderheim were similar to Allegan and Abee in

that the fine fraction is enriched in oldhamite MgS, these results would tend to support the above hypothesis.²

If our identification of the host mineral is valid, one may conclude that the iodine must have been subjected to rather extensive chemical processing after its incorporation in the meteoritic parent bodies. Oldhamite could not have been present in the primordial nebula as such, since it is unstable in the presence of water (the calcium in the nebula would probably be present as CaO or Ca(OH)₂, whereas the sulfur would be present as H₂S plus very stable sulfides such as FeS [Urey, 1952]). Not until the interiors of the meteoritic parent bodies had been heated sufficiently to drive off H₂O could oldhamite be synthesized. *DuFresne and Anders* [1961] have argued on independent grounds that MgS in Chondrites probably was made by the action of sulfur (or sulfur-rich) vapors on magnesian silicates; a similar origin for oldhamite seems likely. During or after its formation, the oldhamite (or MgS) apparently acted as an efficient scavenger for iodine in its environment.

This process almost certainly would have resulted in the strong fractionation of iodine from any Xe¹²⁹ that had accumulated in the primordial iodine-bearing grains. Accordingly, a direct test of the hypothesis proposed by *Eberhardt and Geiss* [1960] and *Zähringer and Gentner* [1960], that the excess Xe¹²⁹ in chondrites is due

² Of course, we have not ruled out adsorption of iodine on fine particles. However, it appears to us that this possibility is very remote. Iodine could not have been present in the primordial nebula as the element or as hydrogen iodide, but only in the form of metal iodides. Adsorption by van der Waals forces is not important for ionic compounds, but only for atoms and covalent molecules.

fractionations in the primordial nebula rather than in the meteorite parent bodies, may be made. If no significant amounts of excess Xe^{129} are found in water-soluble or fine fractions of appropriate meteorites, then I-Xe decay intervals may well refer to events in the primordial nebula. If the excess Xe^{129} is largely found in fractions known to be iodine-rich, then I-Xe decay intervals probably refer to the cooling of the meteorite parent bodies, as we have assumed previously [Goles and Anders, 1960]. The argument would not be changed in any significant way if lawrencite were shown to be the iodine host mineral, since FeCl_2 , too, would be unstable in the presence of H_2S in the primordial nebula. This reasoning does not apply to carbonaceous chondrites, both because there is some question whether they contain any excess Xe^{129} [Goles and Anders, 1961] and because it is very unlikely that oldhamite and lawrencite are present in these meteorites.)

We note that the iodine contents of Canyon Diablo and Toluca troilite and of Toluca metal are great enough so that a search for excess Xe^{129} in these materials might provide another test of the various hypotheses of the origin of this isotopic anomaly. If it should prove feasible to extend the I-Xe dating method to iron meteorites, it might be possible to deduce a cooling curve, and thence an estimate of thermal gradients and over-all thermal conductivities, for the meteoritic parent bodies.

It should also be noted that, if the earth's cosmic complement of sulfur has in large part been fractionated into the core (as FeS), much of the terrestrial iodine may conceivably have been removed from the crust and mantle by this fractionation. This would introduce several uncertainties into the calculation of an I-Xe decay interval for the earth [Goles and Anders, 1961] (Kuroda, unpublished manuscript; Cameron,

unpublished manuscript), particularly with respect to the proper choice of a terrestrial iodine abundance.

Acknowledgment. This work was supported in part by the U.S. Atomic Energy Commission.

REFERENCES

- Buddhue, J. D., Some soluble constituents of meteorites, *Am. Mineral.*, **26**, 677, 1941.
 DuFresne, E. R., and E. Anders, On the retention of primordial noble gases in the Pesyanoe meteorite, *Geochim. et Cosmochim. Acta*, in press, 1961.
 Eberhardt, P., and J. Geiss, Comment on the age of the elements, *Zeit. für Naturforsch.*, **15a**, 547, 1960.
 Goldschmidt, V. M., *Geochemistry*, Oxford University Press, Oxford, 730 pp., 1954.
 Goles, G. G., and E. Anders, Iodine content of meteorites and their I^{129} - Xe^{129} ages, *J. Geophys. Research*, **65**, 4181-4184, 1960.
 Goles, G. G., and E. Anders, The record in the meteorites, On the chronology of the early solar system, *J. Geophys. Research*, **66**, 889-898, 1961.
 Goles, G. G., R. A. Fish, and E. Anders, The record in the meteorites, 1, The former environment of stone meteorites as deduced from K^{40} - Ar^{40} ages, *Geochim. et Cosmochim. Acta*, **19**, 177, 1960.
 Reed, G. W., K. Kigoshi, and A. Turkevich, Concentrations of some heavy elements in meteorites by activation analysis, *Geochim. et Cosmochim. Acta*, **20**, 122, 1960.
 Reynolds, J. H., Determination of the age of the elements, *Phys. Rev. Letters*, **4**, 8, 1960a.
 Reynolds, J. H., Isotopic composition of xenon from enstatite chondrites, *Zeit. für Naturforsch.*, **15a**, 1112, 1960b.
 Tassin, W., On the occurrence of calcium sulphide (oldhamite) in the Allegan meteorite, *Proc. U. S. Natl. Mus.*, **34**, 433, 1908.
 Urey, H. C., *The Planets*, Yale University Press, New Haven, 1952.
 Zähringer, J., and W. Gentner, Uredelgase in einigen Steinmeteoriten, *Zeit. für Naturforsch.*, **15a**, 600, 1960.
 Zähringer, J., and W. Gentner, Zum Xe^{129} im Meteoriten Abee, *Zeit. für Naturforsch.*, **16a**, 239, 1961.

(Received May 8, 1961; revised June 19, 1961.)

First Pulsed Radio Soundings of the Topside of the Ionosphere

R. W. KNECHT AND T. E. VAN ZANDT

National Bureau of Standards, Boulder, Colorado

S. RUSSELL

Airborne Instruments Laboratory
Deer Park, Long Island, New York

At 1817 75°W time on June 24, 1961, a two-frequency pulsed radio sounder (ionosonde) was carried to an altitude of about 1060 km, east of Wallops Island, Virginia, by means of a four-stage, solid-propellant rocket (NASA 8.15). Reflections from the topside of the ionosphere were obtained on one or both frequencies during 13 minutes of the flight. The objective of the experiment was to test the sounding system that is to be used in a satellite-borne, fixed-frequency topside sounder scheduled to be placed in orbit in 1962. The purpose of this short communica-

tion is to present the topside soundings obtained during the flight and to describe some preliminary results of the experiment. A more complete account will be published in a future issue of the *NBS Journal of Research (Radio Propagation)*. The technique involved was essentially that used by conventional bottomside sounders except that the rocket-borne ionosonde was completely transistorized and operated on two fixed frequencies, 4.07 and 5.97 Mc/s, instead of being swept over a wider range. The transmitted pulse length was 100 μ sec, the pulse repetition

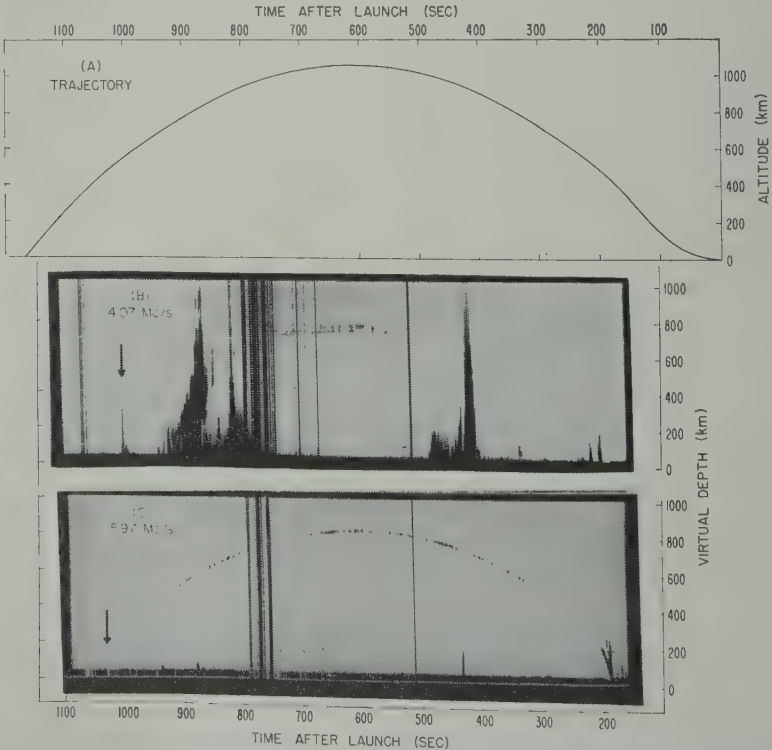


Fig. 1. (A) Trajectory of the flight, 1817 EST, June 24, 1961, Wallops Island, Virginia. (B) 4.07 Mc/s record during the flight. (C) 5.97 Mc/s record during the flight.

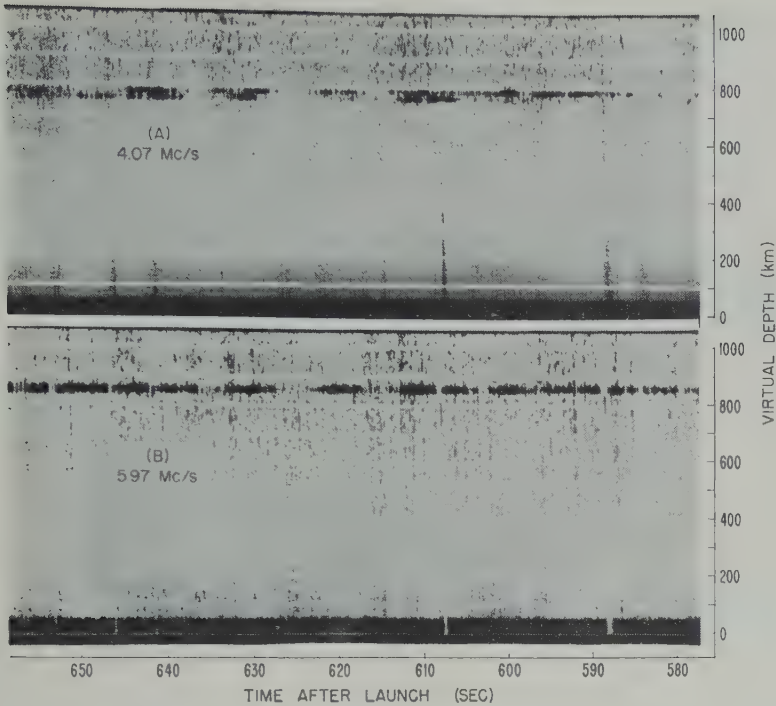


Fig. 2. Topside echoes received on 4.07 Mc/s (A), and 5.97 Mc/s (B) near the apogee of the flight.

was 22 per second on each frequency, and the peak pulse-power radiated on each frequency from the 32-foot dipole antenna was estimated to be 2 to 3 watts. The soundings data were teletyped to several ground observing stations by a frequency-modulation system operating at 240 Mc/s and were recorded on magnetic tape. The sounder was automatically turned on 10 seconds after launch, when the payload was at an altitude of about 360 km. It continued to operate until re-entry into the dense lower atmosphere (80 to 100 km) about 16 minutes later. The trajectory of the flight is shown in Figure 1.

Film records of the 4 and 6 Mc/s data are reproduced in Figures 1(B) and 1(C). The ordinate on the film records is the 'virtual depth' of the reflection level below the rocket sounder, which is obtained in the usual way by multiplying one-half the echo delay time by the speed of light in a vacuum. The virtual depths would be expected to be zero at the times when the sounder emerges from the reflection levels for the frequency and magnetoionic component in question and then increase to a maximum value at the

time when the sounder is farthest from the reflection level, near the apogee of the trajectory. The virtual depths should be symmetrical on the descent. On the best film records the topside echoes conform rather well to the expected pattern. Unfortunately, on the films reproduced in Figure 1, which have a compressed time scale, the weaker 4 Mc/s trace is only barely discernible. Better traces are shown in Figures 2(A) and 2(B), which are sections of expanded film records. The double return on the 4 Mc/s record is due to the presence of two magnetoionic components.

Enlarged views of the topside 'exit' phenomena are shown in Figure 3. On both frequencies unusual effects or 'splashes' (which may be a kind of plasma resonance effect) are observed as the sounder emerges from the topside of the ionosphere and passes through levels corresponding to the three reflection levels [$X = 1$, ordinary (o); $X = 1 - Y$, extraordinary (x); $X = 1 + Y$, (z)]. Also, on the 6 Mc/s x component, and to a much lesser extent on the 4 Mc/s z component, harmonically related traces are observed during the first few seconds after

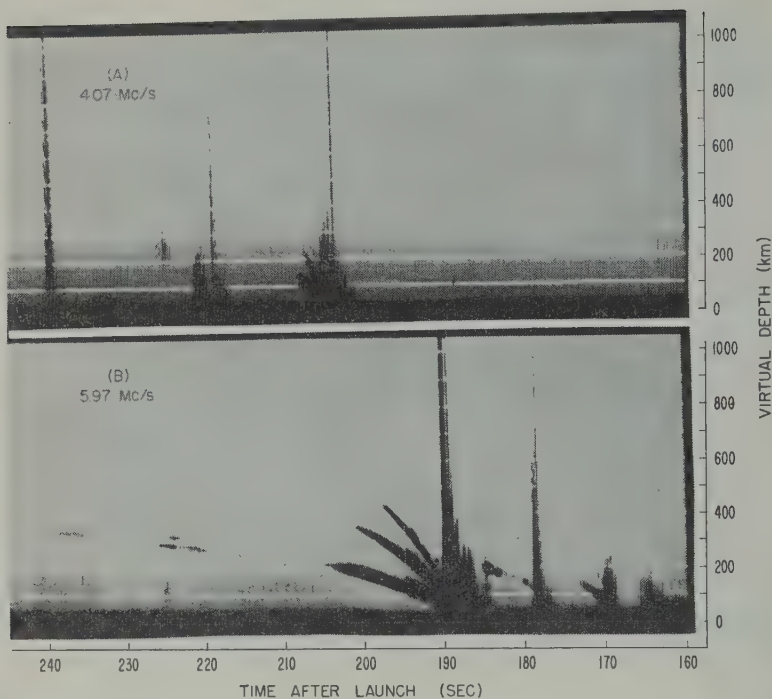


Fig. 3. Topside 'exit' phenomena on 4.07 Mc/s (A) and 5.97 Mc/s (B).

topside exit. These traces appear to be multiple echoes between the sounder and the ionospheric reflection level, which can be understood if the signals propagate in relatively loss-free field-aligned ducts [Gallet and Utlant, 1961].

During the time when the sounder was within about 100 to 150 km of the reflection level, the echoes are weak or missing entirely (Figs. 1 and 3). This is undoubtedly due to detuning of the system by the relatively high electron density in this region. The periodic fluctuations in the strength of the topside echoes and the occasional vertical gaps in the records due to telemetry outages are probably caused by tumbling of the payload.

An increase in background noise, probably cosmic, is evident during the higher parts of the flight.

Prominent effects are observed on the 4 Mc/s record (Fig. 1 (B)) from about 400 to 480 seconds (910 to 1000 km) and from about 780 to 940 seconds (960 to 700 km). Particularly large effects occurred in the 910 to 950 km region on both the ascent and descent. However, they lasted over a much wider height range on the descent, and a second large effect occurred as

the sounder passed through the 840-km level. Similar, but much smaller, effects were observed on the 6 Mc/s record during the same period. The causes of these effects are not understood but it seems likely that they indicate the presence of scattering irregularities in the 700-1000-km height range.

Shortly after the topside echoes cease, owing to detuning on the descent of the flight, splash echoes that are believed to indicate ionospheric re-entries are observed on both frequencies (arrows in Fig. 1 (B) and (C)). Why the three magnetoionic components are not observed individually, as in the topside exit, is not clear. After an interval during which the payload was passing through the over-dense region, it breaks out of the bottomside of the *F'* layer with sharp increases in interference observed almost simultaneously on both frequencies. Prior to bottomside exit, the ionosphere, of course, shielded the sounder from interfering terrestrial signals. The relatively large time difference between the 4 and 6 Mc/s topside re-entries (about 36 seconds) together with the almost simultaneous bottomside exit (about 7 seconds difference) demonstrates a substantial difference between the topside and

topside electron density gradients. On the basis of the exit heights alone, the topside electron-density distribution was nearly exponential, and the neutral atmosphere scale height between 400 and 600 km was about 75 km at the time of the experiment. At the time of the firing, the layer penetration frequency (f_oF_2) was 7.8 MHz/s, and the height of the peak electron density in the F_2 layer was about 305 km, as determined by the Wallops Island bottomside radar.

Acknowledgment. This experiment was a joint

undertaking by the National Bureau of Standards (CRPL) and the Airborne Instruments Laboratory (a Division of Cutler-Hammer, Inc.) with support and technical guidance from the National Aeronautics and Space Administration (Goddard Space Flight Center).

REFERENCE

Gallet, R. M., and W. F. Utlant, Evidence on the laminar nature of the exosphere obtained by means of guided high-frequency wave propagation, *Phys. Rev. Letters*, 6, 591-595, 1961.

(Received August 14, 1961.)

Corrigenda

Dr. I. M. Longman has called attention to an error in his paper 'Formulas for Computing the Tidal Accelerations due to the Moon and the Sun,' which was published in the December 1959 issue of the *Journal*.

Equation 6 should be changed to a vector equation.

$$\mathbf{h}_0 = \mathbf{h}_m + \mathbf{h}_s$$

where the vectors \mathbf{h}_m , \mathbf{h}_s make angles ψ_m , ψ_s , respectively, with the north direction. These angles are given by

$$\cos \psi_m = (\sin l \sin I$$

$$- \sin \lambda \cos \theta) / (\cos \lambda \sin$$

$$\cos \psi_s = (\sin l_1 \sin \omega$$

$$- \sin \lambda \cos \varphi) / (\cos \lambda \sin$$

\mathbf{h}_m is directed toward the moon and so is east or west of north according as l is east or west of north. Similarly \mathbf{h}_s is so governed by l_1 .

Equation 4 for the magnitude h_s of \mathbf{h}_s should read

$$h_s = \frac{3}{2} \frac{\mu S r}{D^3} \sin 2\varphi$$

Dr. S. V. Venkateswaran has called attention to the following typographical errors in the manuscript entitled 'Determination of the Vertical Distribution of Ozone by Satellite Photometry,' by S. V. Venkateswaran, J. G. Moore, and A. J. Krueger, which appeared in June 1961 issue of this journal:

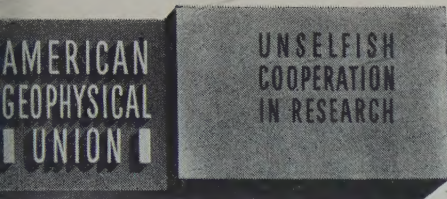
1. a'_{0j} in equations 6 and 8 should read as a'_{ij} and $\sec \rho_i$ in equations 17 and 23 as $\sec \rho_0$.

2. $(\Delta h)_i$ in equations 20 and 21 should read Δh_{ii} .

3. $\sum_{k=j+1}^{k=i} a_{ik}$ in (3) should be

$$\frac{1}{2} \sum_{k=j+1}^{k=i} a_{ik}.$$

These errors have not affected any result or conclusion of the paper.



AMERICAN GEOPHYSICAL UNION

1515 Massachusetts Avenue, N.W., Washington 5, D. C.

Established by the National Research Council in 1919 for the development of the science of geophysics through scientific publication and the advancement of professional ideals.

QUALIFICATIONS FOR MEMBERSHIP

The Membership of the AGU shall consist of Members, Associate Members, Student Members, Corporation Members, and Supporting Members.

Those eligible as candidates for election to the grade of MEMBER shall be:

MEMBER (a) Persons who have made an active contribution to geophysical research through observation, publication, teaching, or administration. Definite evidence should be presented to the Membership Committee. "Publication" may include books, articles, unpublished manuscripts, inventions, or development of geophysical instruments.

(b) Persons who have made active practical application of geophysical research. It should be shown that the nominee's work has not been purely routine, but that it has tended to create new knowledge of, or to broaden or strengthen the application of, geophysical research. In general, the minimum qualifications for membership will be not less than three years of professional experience in some phase of geophysics.

(Continued on next page)

Cut along this line

APPLICATION FOR MEMBERSHIP

Please refer to qualifications on reverse side and designate below type of membership desired:

Member (\$10) ☐ Associate (\$10) ☐ Student (\$4.50) ☐
(1961)

Application forms for Corporation and Supporting Membership are available upon request.

1.	Surname	First Name	Middle Name			
2.	Preferred mailing address for publications					
	Permanent address					
3.	Place	Month	Day	Year of Birth	4.	Country of citizenship/naturalization
5.	Nature of work and title and/or military rank; name and address of organization with which you are associated.					

3. Check section or sections with which affiliation is desired.

- | | |
|--|---|
| <input type="checkbox"/> Geodesy | <input type="checkbox"/> Oceanography |
| <input type="checkbox"/> Seismology | <input type="checkbox"/> Volcanology, Geochemistry, and Petrology |
| <input type="checkbox"/> Meteorology | <input type="checkbox"/> Hydrology |
| <input type="checkbox"/> Geomagnetism and Aeronomy | <input type="checkbox"/> Tectonophysics |

EXPERIENCE (List below, use added sheets as necessary)

Dates: From To Name and address of organization Title, duties, nature of work

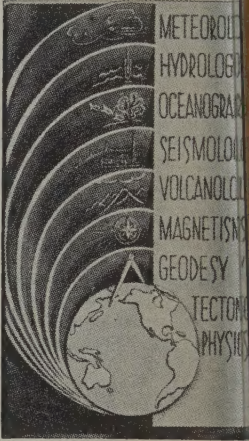
EDUCATION (List Below, use added sheets as necessary)

Dates: From To School Address Major Subject Degree, if any; year
(over)

(Continued from previous page)

Those eligible as candidates for election to the grade of ASSOCIATE MEMBER shall be:

- ASSOCIATE MEMBER Persons who have an active interest in physical processes of the Earth or technical assistance in the application of geophysics. In general, the minimum qualifications for associate membership will be acceptable training or experience in some field of geophysics or allied science.
- CORPORATION AND SUPPORTING MEMBER Corporations and other interested organizations shall be eligible as candidates for election to CORPORATION or SUPPORTING MEMBERSHIP. They shall have the privilege of designating a representative who has the rights and privileges of Members (use special form).
- STUDENT MEMBER Those eligible as candidates for election to the grade of STUDENT MEMBER shall be persons who are graduate or undergraduate students in residence at least half-time and who are specializing in the geophysical sciences. Teaching or research assistants enrolled in more than half of a full-time academic program may also be eligible for Student Membership. Student Members shall have all the privileges of Members except that they shall not vote or hold office.



Cut along this line

- *9. References: Please list below names and addresses of two or three references; include members of the AGU or others who know you well.
- *10. Titles of technical contributions or publications, particularly those in the geophysical sciences, and where published.
- *11. Brief statement of any special interests or qualifications in the geophysical sciences.

Date _____
Written Signature _____

12. (STUDENT MEMBERS ONLY) The person whose signature appears above is known to me and is a student majoring in _____ (subject) at _____

(Name of college or university) expected to graduate in _____ (year) with the degree of _____

☐ He is a full-time student, or ☐ a teaching or research assistant enrolled in more than half of full-time academic program.

(Signature of faculty sponsor) ☐ Check here if faculty sponsor is a member of AGU and willing to act as a regular sponsor for associate membership as well.

(Typed or printed name of sponsor) (Title) _____

* Applicants for student membership may omit Questions 9, 10, and 11, but must fill in Question 12. Please return form with check or money order payable to American Geophysical Union, 1515 Massachusetts Ave., N.W., Washington 5, D. C.

Contents

(Continued from back cover)

ote on the Turbulence Generated by Gravity Waves.....	<i>O. M. Phillips</i>	2889
g-Period Surface Waves from the Chilean Earthquake of May 22, 1960, Recorded on Linear train Seismographs.....	<i>James N. Brune, Hugo Benioff, and Maurice Ewing</i>	2895
asurement of <i>Q</i> for Very Long Period Free Oscillations <i>Leonard E. Alsop, George H. Sutton, and Maurice Ewing</i>		2911
mic Profiler.....	<i>J. I. Ewing and G. B. Tirey</i>	2917
Effect of Plasticity on Decoupling of Underground Explosions <i>A. L. Latter, E. A. Martinelli, J. Mathews, and W. G. McMillan</i>		2929
atic Theory of the Seismic Coupling of a Contained Underground Explosion <i>Norman A. Haskell</i>		2937
Seismic Head Pulse, Reflection and Pseudo-Reflection Pulses.....	<i>C. H. Dix</i>	2945
stic Wave Propagation in Layered Anisotropic Media.....	<i>Don L. Anderson</i>	2953
ase and Group Velocites of Rayleigh Waves in a Spherical, Gravitating Earth <i>Bruce A. Bolt and James Dorman</i>		2965
ady-State Calculations Bearing on Geological Implications of a Phase-Transition Mohorovicic Discontinuity.....	<i>G. W. Wetherill</i>	2983
Use of Hornblendes and Pyroxenes for K-Ar Dating.....	<i>Stanley R. Hart</i>	2995
ural Variations in Isotopic Abundances of Silicon.....	<i>David Tilles</i>	3003
iations of Silicon Isotope Ratios in a Zoned Pegmatite.....	<i>David Tilles</i>	3015
netic Properties of Diabase Dikes.....	<i>D. W. Strangway</i>	3021
Behavior of Volatiles on the Lunar Surface <i>Kenneth Watson, Bruce C. Murray, and Harrison Brown</i>		3033
magnetic and Solar Data.....	<i>J. Virginia Lincoln</i>	3047
eters to the Editor:		
Radiation Measurements in the Slot between the Van Allen Belts to an Altitude of 1415 Kilometers.....	<i>C. A. Gurtler</i>	3050
Rocket Measurement of a Daytime Electron-Density Profile up to 620 Kilometers <i>J. E. Jackson and S. J. Bauer</i>		3055
The Total Electron Content through the Entire Ionosphere by the Faraday Rotation Technique <i>O. Burkard</i>		3058
Some Relationships between Short-Wave Fadeouts, Magnetic Crochets, and Solar Flares <i>Loren W. Acton</i>		3060
Airglow and the South Atlantic Geomagnetic Anomaly.....	<i>K. D. Cole</i>	3064
Possible Effects of Magnetic Field Alignment on Meteoric Ionization.....	<i>Walter F. Bain</i>	3065
A Comparison of Theory and Observation of the Echo I Satellite.....	<i>R. Bryant</i>	3066
Recalculations of Cloud Electrification Based on a General Charge-Separation Mechanism <i>J. D. Sartor</i>		3070
Comment on Paper by B. Le Méhauté, 'A Theoretical Study of Waves Breaking at an Angle with a Shore Line'.....	<i>P. Groen</i>	3072
Author's Reply to the Preceding Comment.....	<i>B. Le Méhauté</i>	3072
A Note on Ocean Sediment Thickness from Surface Wave Dispersion <i>Robert Kovach and Frank Press</i>		3073
On the Geochemical Character of Iodine in Meteorites....	<i>Gordon C. Goles and Edward Anders</i>	3075
First Pulsed Radio Soundings of the Topside of the Ionosphere <i>R. W. Knecht, T. E. Van Zandt, and S. Russell</i>		3078
igendum.....	<i>I. M. Longman</i>	3082
igendum.....	<i>S. V. Venkateswaran, J. G. Moore, and A. J. Krueger</i>	3082

Contents

Dynamics and Structure of the Outer Radiation Belt.....	<i>C. Y. Fan, P. Meyer, and J. A. Simpson</i>	21
The Distribution of Trapped Particles in a Changing Magnetic Field.....	<i>E. N. Parker</i>	22
Forbush Decrease of the Flux of Heavy Primary Nuclei of Cosmic Rays on May 12 and July 12, 1959	<i>S. Biswas</i>	22
Radioactivity Produced in Discoverer XVII by November 12, 1960, Solar Flare Protons	<i>John T. Wasson</i>	22
The Lunar Neutron Flux.....	<i>R. E. Lingenfelter, E. H. Canfield, and W. N. Hess</i>	23
Transresonant Electron Acceleration.....	<i>E. N. Parker</i>	23
A Note on Whistlers Observed above the Ionosphere	<i>Joseph C. Cain, I. R. Shapiro, J. D. Stolarik, and J. P. Heppner</i>	23
Drift Rate in a Dipole Field.....	<i>John S. Lew</i>	23
Rocket Measurements of the Magnetic Field above New Mexico....	<i>R. Hutchinson and B. Shuman</i>	23
Auroral Effects on the Earth's Electric Field.....	<i>George D. Freier</i>	23
The Spectrum and the Total Intensity of Electromagnetic Waves Scattered from an Ionized Gas in Thermal Equilibrium in the Presence of a Static Quasi-Uniform Magnetic Field	<i>J. Renau, H. Camnitz, and W. Flood</i>	27
Prediction of VLF Diurnal Phase Changes and Solar Flare Effects.....	<i>W. D. Westfall</i>	27
VLF Ionospheric Reflection Coefficients—Derivation from Impedance Concepts and Values for Some Model Ionospheres.....	<i>Edward C. Field and Paul Tamarkin</i>	27
The Conductive 'Chapman' Layer.....	<i>C. H. Cummach</i>	27
Analysis of the F_1 - F_2 Transition Region.....	<i>G. A. M. King</i>	27
Effect of Electron-Ion Collisions in the F Region of the Ionosphere on the Absorption of Cosmic Radio Noise at 25 Mc/s at Ahmedabad: Changes in Absorption Associated with Magnetic Storms	<i>K. R. Ramanathan, R. V. Bhonsle, and S. S. Degaonkar</i>	27
Suggestion of a Second Isopycnic Level at 80 to 90 Kilometers over Churchill, Canada..	<i>Allen E. Cole</i>	27
A Model of Solar-Flare-Induced Ionization in the D Region.....	<i>R. C. Whitten and I. G. Poppoff</i>	27
Terrestrial Extremely Low Frequency Noise Spectrum in the Presence of Exponential Ionospheric Conductivity Profiles.....	<i>Janis Galejs</i>	27
Micrometeorite Measurements—Satellite and Ground-Level Data Compared.....	<i>W. D. Crozier</i>	27
On the Long-Period Lunar and Solar Effects on the Motion of an Artificial Satellite, 2....	<i>Peter Musen</i>	27
The Aeolosphere and Atmosphere of Venus.....	<i>E. J. Öpik</i>	28
The Meteorological Rocket Network—An Analysis of the First Year of Operation	<i>Joint Scientific Advisory Group to the Meteorological Rocket Network</i>	28
Experimental Flight Verification of the Economical Net Radiometer	<i>R. H. Bushnell and V. E. Suomi</i>	28
A Method of Obtaining Detailed Wind Shear Measurements for Application to Dynamic Response Problems of Missile Systems.....	<i>Harold B. Tolefson and Robert M. Henry</i>	28
Mean Roughness Factor as a Function of Wind Velocity.....	<i>Allen H. Schooley</i>	28
How Dry Is the Sky?.....	<i>Murray Gutnick</i>	28
Lava-Sea-Air Contact Areas as Sources of Sea-Salt Particles in the Atmosphere	<i>A. H. Woodcock and A. T. Spencer</i>	28

(Continued inside back cover)